Hilary Putnam
Las mil caras del realismo

Introducción de
Miguel Ángel Quintanilla

Paidós I.C.E. | U.A.B.

Las mil caras del realismo Hilary Putnam

B29
P8718

Putnam

31

Hilary Putnam
Las mil caras del realismo

Desde que Kant estableciera la distinción entre la filosofía académica o escolástica, y la filosofía mundana o universal, ha persistido la dificultad de tender un puente entre ambas, esto es, de ser capaz de usar los conceptos y los argumentos técnicos del filósofo de tal manera que sirvan para iluminar los problemas filosóficos que interesan al hombre de la calle, y, a la inversa, de introducir en la discusión técnica de los filósofos cuestiones relevantes para la vida humana. Pues bien, ésta es una de las características del trabajo filosófico de Hilary Putnam, especialmente apreciable en *Las mil caras del realismo*.

Putnam es uno de los filósofos americanos más importantes de nuestros días. Entre sus muchas aportaciones a la discusión filosófica hay una por la que en los últimos años se ha hecho más famoso: su teoría del *realismo interno*. Con ella pretende dar respuesta a una de las preguntas más inquietantes de la filosofía: ¿hasta qué punto podemos decir, con cierto fundamento, que conocemos el mundo real? Lo que nos ofrece en esta pequeña obra maestra es una síntesis de las consecuencias de esa teoría cuando el mismo tipo de argumentación que la sustenta se aplica al resto de las grandes cuestiones filosóficas, en especial a la de la filosofía moral.

Miguel A. Quintanilla, autor de la introducción, es catedrático de Lógica y Filosofía de la Ciencia en la Universidad de Salamanca y autor de numerosos libros y ensayos.

ISBN 84-7509-980-7

46031

9 788475 099804

PAIDÓS

Diseño: Mario Eskenazi

Los contenidos de este libro pueden ser reproducidos en todo o en parte, siempre y cuando se cite la fuente y se haga con fines académicos y no comerciales

Las mil caras del realismo

PENSAMIENTO CONTEMPORÁNEO
Colección dirigida por Manuel Cruz

1. L. Wittgenstein, *Conferencia sobre ética*
2. J. Derrida, *La desconstrucción en las fronteras de la filosofía*
3. P. K. Feyerabend, *Límites de la ciencia*
4. J. F. Lyotard, *¿Por qué filosofar?*
5. A.C. Danto, *Historia y narración*
6. T. S. Kuhn, *¿Qué son las revoluciones científicas? y otros ensayos*
7. M. Foucault, *Tecnologías del yo*
8. N. Luhmann, *Sociedad y sistema: la ambición de la teoría*
9. J. Rawls, *Sobre las libertades*
10. G. Vattimo, *La sociedad transparente*
11. R. Rorty, *El giro lingüístico*
12. G. Colli, *El libro de nuestra crisis*
13. K. O. Apel, *Teoría de la verdad y ética del discurso*
14. J. Elster, *Domar la suerte*
15. H. G. Gadamer, *La actualidad de lo bello*
16. G. E. M. Anscombe, *Intención*
17. J. Habermas, *Escritos sobre moralidad y eticidad*
18. T. W. Adorno, *Actualidad de la filosofía*
19. T. Negri, *Fin de siglo*
20. D. Davidson, *Mente, mundo y acción*
21. E. Husserl, *Invitación a la fenomenología*
22. L. Wittgenstein, *Lecciones y conversaciones sobre estética, psicología y creencia religiosa*
23. R. Carnap, *Autobiografía intelectual*
24. N. Bobbio, *Igualdad y libertad*
25. G. E. Moore, *Ensayos éticos*
26. E. Levinas, *El Tiempo y el Otro*
27. W. Benjamin, *La metafísica de la juventud*
28. E. Jünger y M. Heidegger, *Acerca del nihilismo*
29. R. Dworkin, *Ética privada e igualitarismo político*
30. C. Taylor, *La ética de la autenticidad*
31. H. Putnam, *Las mil caras del realismo*
32. M. Blanchot, *El paso (no) más allá*
33. P. Winch, *Comprender una sociedad primitiva*

Hilary Putnam

Las mil caras del realismo

Introducción de Miguel Ángel Quintanilla

PAIDÓS

Ediciones Paidós
I.C.E. de la Universidad Autónoma de Barcelona
Barcelona - Buenos Aires - México

Título original: *The many faces of realism*
Publicado en inglés por Open Court Publishing Company

Traducción de Margarita Vázquez Campos y
Antonio Manuel Liz Gutiérrez

Cubierta de Mario Eskenazi

B29
P8718
m/766595

DE
MEXICO

FILOSOFIA
Y LETRAS

1.ª edición, 1994

© 1987 by Open Court Publishing Company, La Salle, Ltd.
© de todas las ediciones en castellano,
Ediciones Paidós Ibérica, S.A.,
Mariano Cubí, 92 - 08021 Barcelona
y Editorial Paidós, SAICF,
Defensa, 599 - Buenos Aires

ISBN: 84-7509-980-7
Depósito legal: B-35.088/1994

Impreso en Novagràfik, S.L.,
Puigcerdà, 127 - 08019 Barcelona

Impreso en España - Printed in Spain

A la memoria de mi querido tío

PETER SAMPSON

cuya ayuda y estímulo hicieron posible, hace muchos años, mi educación como graduado

F253321

SUMARIO

Prólogo del autor a la edición española .. 11

Introducción: El realismo necesario, por *Miguel A. Quintanilla* 17

LAS MIL CARAS DEL REALISMO

Prefacio 39

Conferencia I: ¿Queda todavía algo por decir acerca de la realidad y la verdad? .. 41

Propiedades intrínsecas: las disposiciones .. 48
Propiedades intrínsecas: la intencionalidad . 52
¿Por qué la intencionalidad es un problema tan insoluble? 55
«El rastro de la serpiente humana está por por todas partes» 60

Conferencia II: El realismo y la razonabilidad 69

Realidad sin dicotomías 81

Conferencia III: La igualdad y nuestra imagen moral del mundo 93

El intento de la Escuela de Francfort por justificar la igualdad 110

Democracia sin una imagen moral 116

Conferencia IV: La razonabilidad como hecho y como valor 125

El método científico 139
El «problema epistemológico» 145
La importancia del puzzle de Peirce 153

PRÓLOGO DEL AUTOR A LA EDICIÓN ESPAÑOLA

Estoy especialmente contento de que *Las mil caras del realismo* esté ahora disponible para los lectores de habla hispana, porque es, de mis libros, con el que más me gustaría que empezara un lector. De hecho, es mi favorito entre mis primeros libros, porque en él presento lo que veo todavía como mi propia «imagen moral», si no del mundo, al menos del lugar de la filosofía en el mundo. Intento también mostrar cómo algunos problemas de disciplinas filosóficas que son tratadas de manera convencional como diferentes, con nombres como «metafísica», «ética», «epistemología», «filosofía de la lógica» y «filosofía del lenguaje», están íntimamente conectados y cómo nos enfrentamos a la misma elección fundamental con respecto a todos estos problemas. Esa elección, como explico en las conferencias que siguen, es la elección entre tomar en serio los conceptos (y las acciones y formas de vida con las que esos conceptos están entrelazados) que encontramos realmente indispensables tanto en nuestra vida práctica como teórica y buscar un principio metafísico que nos diga cuáles de esos conceptos describen la realidad «como es en sí misma» y cuáles son «proyecciones humanas». La primera alternativa es la alternativa que yo denomino «realismo pragmático» o «realismo interno» en estas conferencias, y la conecto con una idea que heredé de William James, la idea de tomar en serio el

punto de vista del agente cuando hacemos filosofía. Es una característica de la segunda alternativa, la alternativa filosófica más común, el no ver siquiera la primera alternativa, el punto de vista del agente, como algo que se ha de tomar en serio a la hora de hacer filosofía. En vez de ello, los filósofos que eligen la segunda alternativa ven las opciones de una manera bastante diferente. Todos y cada uno de ellos caen en la falacia de pensar que sabemos cuáles son las opciones filosóficas y que ellas equivalen en cada caso a una elección obligada entre un curioso algo metafísico situado detrás de nuestro discurso (tanto si este discurso es sobre lo correcto y lo erróneo, sobre lo objetivo y lo subjetivo, sobre la causalidad o sobre el tema filosófico que sea) y un reduccionismo «fuertemente mentalista». Pero, como intento mostrar en estas conferencias, los reduccionismos en apariencia «fuertemente mentalistas» son incoherentes, mientras que los curiosos algos metafísicos son ininteligibles.

Mi rechazo a la idea de dividir las cosas y las propiedades de las que hablamos en aquellas que están «nouménicamente ahí» y aquellas que son sólo «proyecciones» puede traer a la memoria de algunos lectores los puntos de vista de Richard Rorty, un filósofo del que discrepo con respeto en estas páginas (así como de otros filósofos franceses contemporáneos). Rorty dice también que no podemos hacer ese trabajo dicotómico. Pero hay una diferencia fundamental entre el sentido en el que él rechaza ésta y las otras dicotomías que yo critico en las páginas que siguen y el sentido en que yo lo hago. Por ejemplo, si el tipo de realismo metafísico que postula «cosas en sí mismas» con una «naturaleza intrínseca» no tiene sentido, entonces, concluye

Rorty, tampoco lo tiene la noción de *objetividad*. Deberíamos suprimir todo el discurso sobre objetividad y, en su lugar, hablar de «solidaridad».[1] Deberíamos convertirnos todos en relativistas culturales.

Éste es un punto de vista que encuentro catastrófico.[2] Aunque me uno a Rorty en el rechazo de varias nociones metafísicas tradicionales, no creo que abandonar estas nociones (por encontrarlas, finalmente, vacías) nos obligue a abrazar el relativismo o el irracionalismo. En lo que estoy de acuerdo con Rorty (o con Derrida) es en que la idea de comparar, por un lado, mis pensamientos y creencias con, por otro lado, las cosas «tal como son en sí mismas» no tiene sentido; pero no estoy de acuerdo en que esta idea sea un presupuesto necesario de la idea común de que hay objetos, «animales, vegetales y minerales», que no son partes del pensamiento o del lenguaje o de la idea igualmente común de que lo que nosotros decimos acerca de estos objetos algunas veces *capta los hechos correctamente*.

El diagnóstico que he sugerido en otra parte sobre la difícil situación de Rorty es que él está tan preocupado por la falta de una *garantía* de que nuestras palabras representen las cosas fuera de sí mismas que, al encontrar que una garantía del único tipo que él prevé es «imposible», siente que no tiene otra alternativa que rechazar la propia idea de representación como una equivocación.[3] (Aquí ve-

1. Véase «Solidarity or Objectivity», en Rorty, R., *Objectivity, Relativism and Truth: Philosophical Papers, volume 1*, Cambridge, Cambridge University Press, 1991.

2. Lo analizo con más detalle en «The Question of Realism» un artículo que está incluido en mi libro *Words and Life*, Cambridge, Mass., Harvard University Press, 1994.

3. Véase la nota 2.

mos también cómo Rorty es un realista metafísico decepcionado.) El problema, me parece, es que Rorty ha dejado sin explorar el tipo de «imposibilidad» que está en cuestión cuando concluye que tal garantía es imposible. Rorty tiene razón cuando dice que no tiene sentido pensar en colocarse fuera de los propios pensamientos y conceptos y comparar la «realidad como es en sí misma» con esos pensamientos y conceptos. ¿Cómo podría tener sentido *esa* idea? Lo que Rorty ha hecho es moverse desde la ininteligibilidad de este tipo de garantía hasta un escepticismo acerca de la posibilidad de representación en un sentido perfectamente común de «representación». Él nos deja con la conclusión de que no hay ninguna manera metafísicamente inocente de decir que nuestras palabras representan las cosas fuera de sí mismas. Al haber dejado sin investigar el carácter de la ininteligibilidad que vicia el tipo de realismo metafísico que él ataca, Rorty permanece ciego al modo en que su propio rechazo del realismo metafísico comparte la *misma* ininteligibilidad. Digo «comparte la *misma* ininteligibilidad»; si es ininteligible decir que algunas veces *tenemos éxito* al representar las cosas tal como son en sí mismas, entonces es igualmente ininteligible decir que *nunca tenemos éxito* al representar las cosas tal como son en sí mismas. El modo en el que el escepticismo es el reverso del ansia por un tipo ininteligible de certeza queda ilustrado por la complacencia de Rorty al abandonar el hecho perfectamente obvio de que el lenguaje puede representar algo que está fuera del lenguaje. La verdadera misión de la filosofía aquí es iluminar la noción ordinaria de representación (y de un mundo de cosas que se han de representar); no permanecer inermes

en un gesto de rechazo que está tan vacío como aquello que rechaza.[4]

Los lectores familiarizados con otros libros y artículos en los que uso la expresión «realismo interno» sabrán que en esos escritos[5] relaciono el término con un análisis particular de la noción de verdad; uno que *no* es mencionado en este pequeño libro. Puesto que desde entonces he abandonado ese análisis como erróneo, sin abandonar nada de lo que sostengo en este libro, el hecho de que me refrenase de defender ese análisis erróneo aquí es, tal como lo veo ahora, ¡otra virtud más del presente trabajo! Volviéndolo a examinar después de casi diez años, me parece, de hecho, que este libro anticipó muy bien la dirección en la que mi pensamiento iba a ir en el futuro, a pesar de que yo no lo sabía en ese momento.

HILARY PUTNAM

4. De nuevo analizo cómo puede realizarse esta tarea de clarificación en mis Conferencias Dewey, «Sense, Nonsense, and the Senses: An Inquiry Into the Powers of the Human Mind», *The Journal of Philosophy*, XCI (septiembre, 1994).

5. Por ejemplo, *Reason, Truth and History*, Cambridge, Cambridge University Press, 1981, y *Representation and Reality*, Cambridge, Mass., MIT Press, 1988.

me
de
Ro
da
ga
dic
de
rar
sar
esa
la i
esc
tac
pre
no
de
sas
gar
tip
ma
rea
bili
dac
nen
sí n
cir
tal
cep
teli
cer
me
alg
sió
nar
que

INTRODUCCIÓN

EL REALISMO NECESARIO

Para un lector profano, la mayoría de las grandes obras filosóficas resultan generalmente abstrusas: no sólo dicen con frecuencia cosas complicadas usando un lenguaje difícilmente comprensible, sino que además parecen plantearse problemas que sólo los propios filósofos consideran importantes. Uno de éstos podría ser el siguiente: ¿es *real* el mundo en el que vivimos? Otro: ¿se pueden *justificar* racionalmente las normas morales de conducta? Otro: ¿por qué es conveniente (quizás deberíamos decir «racional») *actuar racionalmente*? Más: ¿qué estamos diciendo realmente cuando decimos que un enunciado *es verdadero*?

En la mayoría de los casos las palabras que usamos para formular estas preguntas son palabras de uso común, que utilizamos en la vida diaria. A veces incluso nos hacemos, también en la vida diaria, preguntas parecidas, aunque nunca solemos llevar la indagación demasiado lejos, en todo caso no tan lejos como para llegar a un punto en el que empezaríamos a dudar de que la pregunta tenga el mismo sentido que en principio suponíamos.

Pues bien, lo que caracteriza al pensamiento filosófico es precisamente que se ocupa de ese tipo de problemas que rara vez nos planteamos en la vida cotidiana, pero que, cuando por alguna razón nos enfrentamos a ellos, terminamos descubriendo lo difícil que resulta llegar a conclusiones inteligibles.

Los filósofos llevan más de veinticuatro siglos discutiendo estos temas y, al cabo de tanto tiempo, han ido construyendo una jerga propia que les permite refinar sus conceptos y sus teorías, hasta un punto que las hace difícilmente accesibles a lectores cultos pero no especializados.

La gran virtud del libro que tenemos entre manos es precisamente ésta: se trata de una obra filosófica en sentido estricto, donde sólo se habla de ese tipo de problemas, en principio abstrusos y aparentemente irreales, que además son discutidos en los mismos términos en que suelen discutirlos los filósofos profesionales. Y sin embargo el libro de Putnam no sólo consigue hacer comprensible la discusión filosófica de esos problemas, sino que sobre todo hace que el lector culto comprenda por qué esos problemas abstrusos son significativos e importantes.

Putnam es uno de los filósofos americanos más importantes de nuestros días. Entre sus muchas aportaciones a la discusión filosófica hay una por la que en los últimos años se ha hecho más famoso: su teoría del realismo interno. Y lo que nos ofrece en esta pequeña obra es una síntesis de las consecuencias de esa teoría cuando se aplica el mismo tipo de argumentación que la sustenta al resto de las grandes cuestiones filosóficas, en especial en el campo de la filosofía moral.

Lo que Putnam llama realismo interno, o realismo pragmático, es desde luego una de esas teorías filosóficas con aires de excesiva complejidad a las que nos referíamos hace un momento. En esencia se trata de dar respuesta a la artificiosa pregunta siguiente: ¿hasta qué punto podemos decir, con cierto fundamento, que conocemos el mundo real?

En esta breve presentación de la obra de Putnam pretendo dos cosas: primero explicar por qué esa pregunta es importante; segundo discutir algunos puntos de la respuesta que nos ofrece Putnam.

***

A poco que nos paremos a pensar, descubriremos que hay razones de sobra para sospechar que aun los conocimientos que consideramos más seguros e incontrovertibles, tienen un frágil fundamento. Supongamos que alguien está dispuesto a no admitir como válidas más que las opiniones e ideas que estén sólidamente basadas en experiencias y observaciones concretas. Apenas podría entender nada de lo que sucede a su alrededor: tendría que prescindir de todos los conocimientos matemáticos, además debería poner en duda que otras personas puedan compartir sus conocimientos y experiencias e incluso tendría que terminar dudando de la consistencia de sus propias observaciones. En efecto, cada día tenemos la experiencia de que sale un sol por la mañana y no ponemos en duda que el sol que sale hoy es el mismo que salió ayer, a pesar de que de esto nunca hemos tenido una experiencia concreta; tampoco disponemos de ningún metro que valga para medir de forma exacta el lado de un cuadrado y su diagonal, pero no por ello dudamos de que podremos medir la diferencia entre la diagonal y el lado con el grado de precisión que en cada caso necesitemos; y aunque no podemos compartir directamente las experiencias del resto de la gente (precisamente porque son sus propias experiencias) ello no nos impide hablar con otros, entender lo que dicen y compartir a veces sus conocimientos y opi-

niones. Tampoco sería más confortable nuestra posición si nos pasáramos al extremo contrario y nos declaráramos dispuestos a admitir sólo aquellas opiniones e ideas que estuvieran sólidamente fundadas en la lógica y las matemáticas: podríamos entonces hilar largos razonamientos, todos ellos verdaderos, pero vacíos de contenido, de manera que no seríamos capaces de concluir nada relevante acerca de nada real; podríamos utilizar los números para sumar, multiplicar, etc., pero no podríamos contar ovejas con ellos, porque las ovejas no son entidades matemáticas.

Los antiguos filósofos griegos ya se dieron cuenta de todos estos problemas. Y en buena parte la historia de la filosofía, desde hace más de dos mil años, está plagada de respuestas ingeniosas para resolverlos. Platón inventó el mito de la caverna para tratar de explicar cómo era posible el conocimiento del mundo por parte de los hombres. Aristóteles supuso que la facultad humana del entendimiento podía captar las ideas generales a partir de las cosas concretas. Con el desarrollo de la ciencia moderna, el problema volvió a replantearse con mayor crudeza. Muchas de las evidencias del sentido común aceptadas hasta entonces se vinieron abajo; y empezó a aceptarse, como dice Putnam, que «el rastro de la serpiente humana está por todas partes»: es decir que el sujeto humano deja su propia impronta en todas las ideas, representaciones y teorías con las que pretendemos conocer la realidad. De modo que no es posible seguir hablando de un mundo real, independiente de nosotros, y un sujeto humano capaz de conocerlo objetivamente. Kant fue sin duda el filósofo moderno que más claramente definió la situación: el mundo objetivo, tal como lo

conocemos, es el resultado de nuestra propia actividad intelectiva. Con otras palabras: somos nosotros, con nuestra capacidad sensorial e intelectual, quienes organizamos la experiencia en categorías que nos la hacen comprensible y así construimos las cosas y los hechos, las relaciones de causalidad, las regularidades que observamos en la naturaleza, etc.

La diferencia más importante entre Kant y otros filósofos anteriores no reside, sin embargo, en la importancia que se concede a la actividad del sujeto humano en el conocimiento del mundo, sino en la forma como se concibe esa actividad. De acuerdo con Kant el cometido de la razón humana no consiste en descubrir o representar el mundo real sino en construir los conceptos que nos permiten entenderlo.

Hay, sin embargo, en la respuesta kantiana al problema del conocimiento, un aspecto que parece insatisfactorio a simple vista. Por una parte se supone que hay criterios objetivos para distinguir entre opiniones correctas e incorrectas sobre el mundo que nos rodea. Por otra parte debemos admitir que nuestra propia noción de «mundo que nos rodea» depende de nosotros mismos. Con otras palabras: desearíamos poder decir que nuestros conceptos y representaciones se corresponden de alguna manera con el mundo real, pero al mismo tiempo tenemos que admitir que lo que llamamos el mundo real no es nada que podamos identificar al margen de nuestros conceptos y representaciones.

Hay dos respuestas clásicas (y en realidad prekantianas) a este problema: la respuesta dogmática y la respuesta escéptica. La respuesta dogmática (que Putnam identifica con lo que él llama el rea-

lismo metafísico) se puede resumir en las siguientes tesis:

1. El mundo existe independientemente de nuestros conceptos y representaciones.

2. Ese mundo objetivo e independiente tiene una estructura determinada; es decir está compuesto de cosas, propiedades, hechos objetivos que existen independientemente de que nosotros los conozcamos o no.

3. Una representación verdadera o correcta del mundo es aquella que se refiere a objetos realmente existentes en el mundo y describe las propiedades que esos objetos realmente tienen.

4. Existe (aunque no la conozcamos *todavía*) una imagen o representación completa del mundo objetivo tal como es en sí mismo (lo que Putnam identifica como «el punto de vista del Ojo de Dios»).

La respuesta escéptica adopta como punto de partida la constatación de que no existe ninguna forma de establecer una correspondencia fiable entre nuestras representaciones y las supuestas propiedades objetivas del mundo independiente de nosotros. A partir de aquí la conclusión más radical es que no podemos saber ni decir nada con sentido acerca de la realidad, todos nuestros discursos son pura palabrería.

Tanto el realismo metafísico como el relativismo escéptico radical son –ya lo hemos dicho– teorías típicamente prekantianas. Y sin embargo siguen estando de plena actualidad. Hay razones para ello. Por una parte, una cierta actitud realista parece imprescindible para poder entender una parte fundamental de nuestra cultura: el aprecio por la verdad y la íntima convicción de que el avance del conocimiento –especialmente del conocimiento científi-

co– tiene algo que ver con una ampliación efectiva de nuestros conocimientos *verdaderos*, es decir con nuestra capacidad de representar adecuadamente una parte cada vez más importante de la realidad. Por otra parte, no es fácil entender las razones para evitar el escepticismo radical, una vez que nos hayamos decidido a aceptar la visión kantiana del conocimiento «objetivo» como inevitablemente «subjetivo». El propio Kant advertía que es un *escándalo para la filosofía y para la razón humana* el que haya que admitirse la existencia de los objetos exteriores sólo a modo de *creencia*, es decir como una opinión que no podemos justificar racionalmente.

Es en este contexto donde hay que situar la teoría del realismo interno de Putnam. Se trataría de realizar un doble movimiento: por una parte recuperar la visión kantiana de la filosofía, radicalizándola y, por otra, de demostrar que el resultado es plenamente satisfactorio, es decir que podemos recuperar lo que podríamos llamar el contenido *sano* del realismo (lo que Putnam suele llamar el realismo con minúscula), evitando tanto el dogmatismo del Realismo metafísico (el realismo con mayúscula) como el nihilismo al que conduce en último término el escepticismo radical.

Es especialmente relevante a este respecto la introducción del tema del realismo interno que hace Putnam en el primer capítulo de este librito: se trata de recuperar el realismo del sentido común, frente a las pretensiones del realismo científico (del realismo metafísico disfrazado de científico) que, a la postre, resulta incompatible con esa noción de realidad que se pretendía fundamentar. Dicho en términos muy simples, que desde luego no hacen justicia a los finos análisis de Putnam, la situación po-

dría describirse así: el realista ingenuo mantiene dos tipos de pretensiones, por una parte considera que el mundo real es el mundo de su experiencia cotidiana, el mundo repleto de sillas, mesas, personas con las que habla, etc.; por otra parte pretende que ese mundo de su experiencia cotidiana es exactamente una parte del mundo real, es decir del mundo tal como existe por sí mismo, independientemente de cualquier experiencia que cualquier sujeto pueda tener de él. Sin embargo tales pretensiones son incompatibles con el realismo científico, pues la ciencia actual, en especial la ciencia física, nos dice que el mundo de nuestra experiencia cotidiana no es real: no sale y se pone el sol, sino que es la tierra la que da vueltas, no hay sillas y mesas sólidas, sino campos y partículas subatómicas o, peor aún, perturbaciones en el continuo espacio-temporal, etc., y ésta es la *verdadera realidad* del mundo del realista ingenuo. De manera que la única forma de mantener la segunda pretensión del realista ingenuo (la de que el mundo existe independientemente de la experiencia que nosotros tenemos de él) es abandonar la primera (la de que nuestra experiencia cotidiana sea una experiencia del mundo eral), pero al alto precio de tener que concluir que, una vez abandonada esta referencia sólida del realismo, lo que tenemos es un «mundo real» plagado de proyecciones humanas, de creaciones de nuestro espíritu, y además compatible con infinitas creaciones diferentes. Es decir, todo, menos la promesa que nos había hecho el realista metafísico disfrazado de científico.

La salida de Putnam consiste en cambiar de dirección en la revisión de los presupuestos del realismo ingenuo: el problema no es tanto su primera pretensión (la de que el mundo que nos rodea de si-

llas, mesas y cubitos de hielo sea real) sino la segunda: la de que sea legítimo decir que hay un mundo real independiente de nuestra experiencia de las sillas, las mesas y los cubitos de hielo (o, para el caso, también de nuestra experiencia científica de los campos de fuerza, los fenómenos cuánticos y las dimensiones matemáticas del espacio-tiempo). Con otras palabras, podemos ser realistas, como lo somos en la vida cotidiana, lo que no podemos es mantener la ilusión de que se puede hablar de un mundo independiente de nuestra experiencia o de «objetos en sí», al margen de los conceptos con los que representamos esos objetos. El resultado es que el realismo (el realismo pragmático de nuestra experiencia cotidiana) va indisolublemente unido a un cierto relativismo conceptual: la realidad es la realidad tal como es concebida, lo real es siempre relativo al aparato conceptual que utilizamos para hablar o representar el mundo real.

Según Putnam este relativismo conceptual del realismo interno o pragmático no tiene por qué conducir a conclusiones escépticas radicales: aunque los hechos sean relativos a los esquemas conceptuales que utilizamos para describirlos, ello no quiere decir que no podamos distinguir, dentro de un determinado esquema conceptual, entre hechos reales y no reales, o paralelamente, entre enunciados verdaderos y falsos acerca de los hechos. Incluso es posible discutir sobre el valor de diferentes esquemas conceptuales en relación con determinados criterios (de carácter metodológico, moral, etc.) que compartan todos ellos.

Algunos críticos de Putnam han señalado que, si se acepta su concepción del realismo interno, lo difícil es explicar la posibilidad del error en nuestros

juicios acerca de la realidad. Pero creo que éste no es un argumento sólido. Podríamos aceptar, por ejemplo, que todas nuestras teorías matemáticas son independientes de cualquier realidad física, es decir dependen solamente de su propio contenido conceptual, y sin embargo eso sería compatible con el hecho de que también en matemáticas podemos cometer errores y descubrir nuevas verdades. Más aún, cualquier ente puramente ficticio de la literatura –como por ejemplo Don Quijote– es, incluso para un realista metafísico, un mero producto de la imaginación; y sin embargo eso no nos autoriza a añadir a esos entes ficticios atributos arbitrarios; un Don Quijote gordo, comilón y pillo sería un Don Quijote falso. Naturalmente, para ser precisos, habría que decir que es falso que Don Quijote sea gordo, comilón y pillo *dentro del contexto conceptual* en el que nos movemos, según el cual Don Quijote es el personaje principal de una obra escrita por Miguel de Cervantes, etc., etc. Si evolucionara nuestro contexto conceptual hasta incluir, por ejemplo, una teoría según la cual Avellaneda y Cervantes eran en realidad la misma persona, tendríamos que cambiar las propiedades que objetivamente podemos atribuir a Don Quijote (aunque desde luego no tanto como para hacer de él un hombre gordo, comilón y pillo, pero todo puede ser cuestión de tiempo y de adecuados cambios conceptuales).

En conjunto, la filosofía de Putnam ofrece una imagen muy atractiva: podemos mantener todas aquellas ideas que nos parecen dignas de nuestra cultura, incluso refinar muchas de ellas, y en especial nuestra propia imagen moral del mundo. Podemos además argumentar racionalmente sobre ellas, seguir comprometidos con nuestro afán de

objetividad, de progreso cognoscitivo sobre el mundo, con nuestro aprecio por la razón humana, por la libertad y por los valores de la igualdad entre los hombres, etc. La buena nueva es que, para seguir manteniendo (y discutiendo y mejorando) todas estas concepciones filosóficas, plenas de sentido, no es preciso que les añadamos el componente dogmático del realismo metafísico para nuestras creencias sobre el mundo, o de la justificación última y absoluta de nuestros valores morales o de nuestra actitud racional. Más aún, si queremos seguir manteniéndolas de forma coherente y plenamente satisfactoria (si queremos seguir fieles al realismo del sentido común, al valor moral de nuestras tradiciones humanistas, a la confianza en la razón humana misma) lo haremos mejor si nos libramos de las adherencias dogmáticas que terminarían proporcionándonos viles sustitutos del contenido más valioso de esas concepciones, al contaminarlas con el dogmatismo o al diluirlas en el relativismo radical.

Desde luego, hay un pequeño precio que pagar por esta lucidez: es el precio de la modestia intelectual. En filosofía, como en la vida real, hay «lechos de roca» más allá de los cuales nos resulta imposible seguir profundizando. Estamos tan implicados en el mundo, que nosotros mismos construimos, que siempre habrá un punto límite para nuestra capacidad de comprensión. La discusión que hace Putnam al final de su libro, acerca de las razones que tenemos para actuar racionalmente, incluso en situaciones propias de la llamada racionalidad instrumental (la racionalidad referida a la elección de los medios más adecuados para un fin dado), es un ejemplo de hasta qué punto puede ser apasionante la aventura intelectual de seguir un problema filo-

sófico hasta dar, por el momento, con el «lecho de roca» de Wittgenstein.

***

Pero no seríamos fieles al espíritu de la filosofía de Putnam si nos contentáramos con recoger su mensaje sin someterlo a discusión. Valgan pues estos últimos párrafos como invitación al lector para que intente también leer a Putnam desde un punto de vista diferente al que él mismo nos sugiere.

Para empezar podemos hacernos algunas preguntas que merecen nuestra atención. La primera es bien simple: si de lo que se trata es de recuperar el realismo del sentido común ¿no deberíamos pensar que también es parte inherente de este tipo de realismo la creencia en que las sillas y las mesas de nuestra experiencia existen realmente en el mundo, independientemente de los esquemas conceptuales que nosotros utilicemos para hacernos una idea de ellas? La respuesta de Putnam sería que cualquier cosa que podamos decir con sentido acerca de las sillas y mesas será siempre relativo a nuestra forma de concebir sillas y mesas y que la hipótesis metafísica de que existen sillas y mesas «en sí mismas» (el punto de vista del Ojo de Dios) no añade nada a lo que sobre ellas podemos decir sin asumir esa hipótesis, es decir que tales sillas y mesas son reales relativamente a nuestros esquemas conceptuales. Pero la cuestión entonces podría ser ¿concebiríamos las sillas y mesas de la misma forma que las concebimos si no tuviéramos la pretensión de que nuestras ideas acerca de las sillas y las mesas se correspondan de alguna manera con lo que las sillas y las mesas son «en sí mismas»? En

caso de que la respuesta a esta pregunta fuera afirmativa, la conclusión más lógica sería pensar que, incluso desde el punto de vista pragmático, la *hipótesis metafísica* del mundo real independiente (de nuestra experiencia) tiene un valor como mínimo equivalente al que tiene la concepción del realismo interno, según la cual el mundo de nuestra experiencia (las sillas y las mesas, para entendernos) es el mundo real (en el sentido del realismo interno, es decir relativamente a nuestros propios esquemas conceptuales). Ahora bien, yo creo que hay razones para pensar que la *pretensión* de representar una realidad independiente de nuestra propia representación es inherente al realismo del sentido común, así como al realismo del conocimiento científico, al menos en el sentido que acabamos de indicar: muchos de nuestros argumentos para aceptar o rechazar una determinada representación del mundo se basan en que consideramos que la representación corresponde o no corresponde de hecho a una supuesta realidad objetiva. Naturalmente que podemos estar equivocados, y de hecho nos equivocamos con frecuencia, respecto a nuestras pretensiones realistas, pero la cuestión no es ésa sino la de si tiene algún sentido la misma pretensión. Yo creo que, planteada la cuestión en estos términos, la respuesta de Putnam debe ser que esa pretensión no tiene ningún sentido y debe ser abandonada. Y sin embargo, es posible que al abandonarla estemos también olvidando otros aspectos del realismo del sentido común que son pragmáticamente interesantes.

Supongamos que lo que estamos discutiendo no es si una determinada concepción del sentido común es compatible con nuestras teorías científicas o si una determinada teoría científica es más com-

patible con los hechos conocidos que otra (discusiones todas estas que tienen pleno sentido dentro del realismo interno). Por el contrario, de lo que se trata es de argumentar si, en general nuestra forma de concebir el mundo, a partir de la tradición de la ciencia moderna (llamémosla concepción A), es preferible a otra forma alternativa, de carácter místico, por ejemplo (llamémosla la concepción B). Supongamos además que la concepción B incorpora la teoría del realismo interno; de hecho una determinada concepción mística del mundo puede afirmar que la suya es la mejor representación de la realidad, asumiendo que la realidad a la que ella se refiere es propiamente la realidad tal como ésta es construida en su propio contexto conceptual. La cuestión es si, en este caso, sería posible argumentar racionalmente a favor de A (la concepción del mundo compatible con la tradición de la ciencia moderna) sin reivindicar que los méritos de A radican precisamente en que, a diferencia de B, A pretende ser una representación objetiva de una realidad cuya existencia se supone independiente tanto de A como de B. Obsérvese que no estoy proponiendo que A lleve razón: simplemente estoy señalando que *pragmáticamente* no podríamos distinguir los méritos de A frente a los de B sin asumir que el valor de A depende precisamente de su reivindicación de un realismo no interno. Desde luego es lógicamente posible que A sea una concepción equivocada, y que B finalmente resulte una alternativa más razonable; lo que no parece verosímil es que podamos describir la verdadera contraposición entre A y B (entre sus diferentes pretensiones como concepciones aceptables del mundo) sin asumir que A incluye la pretensión de representar objetiva-

mente el mundo tal como es. Dicho de otra forma: parece que si aceptamos el realismo interno de Putnam, con todas sus consecuencias, podemos estar perdiendo algo importante de esa concepción del mundo que el propio Putnam aprecia, y que incluye en buena parte los contenidos más aceptables del realismo del sentido común que reivindica.

No es éste el lugar más indicado para revisar todos los argumentos de Putnam. Pero sí puede ser útil indicar al lector algunas pistas por donde proseguir la discusión con nuestro autor. Una posible vía de salida consiste en revisar las tesis del realismo metafísico que Putnam rechaza. Hemos resumido esta doctrina en cuatro tesis. El punto central del realismo metafísico está en la tercera tesis, que es la tesis de la verdad como correspondencia. Las dos primeras son necesarias para que la concepción de la verdad como correspondencia tenga sentido. La cuarta parece una consecuencia necesaria de la tercera. En otros escritos Putnam aduce numerosos argumentos contra cada una de las tesis del realismo metafísico, aunque parece que considera que las cuatro tesis están íntimamente entrelazadas. Ahora bien, podríamos preguntarnos qué quedaría del realismo metafísico si rechazáramos o modificáramos alguna de las cuatro tesis, manteniendo el resto.

De hecho lo que podríamos llamar *realismo tentativo* o hipotético podría caracterizarse por las siguientes tesis:

1. La realidad objetiva existe independientemente de que la conozcamos o no.
2. La realidad objetiva independiente tiene en sí misma algún grado de organización o estructura.

3. El conocimiento humano, y en especial el conocimiento científico, pretende conseguir una representación adecuada de la realidad objetiva.

4. Es posible alcanzar, a través de la investigación científica, representaciones adecuadas, aunque parciales e incompletas, de la realidad objetiva.

A diferencia del realismo metafísico que critica Putnam, esta versión del realismo no incorpora la tesis del Ojo de Dios, ni la teoría ingenua de la verdad como correspondencia o, mejor aún, como copia o reflejo de la realidad. Sin embargo acepta las tesis externalistas implícitas en el realismo del sentido común. Y además acepta como legítima la pretensión de que el conocimiento científico pueda ser una representación adecuada aunque parcial e incompleta del mundo real.

En realidad no hay tanta distancia entre las consecuencias filosóficas de este realismo tentativo y las del realismo interno de Putnam. De hecho, el realismo tentativo puede aceptar, con Putnam, que sólo podemos hablar acerca del mundo real a través de nuestros esquemas conceptuales, pero de ahí se infiere que nuestras representaciones del mundo real son siempre parciales e incompletas, no que carezca de sentido suponer que representan algo independiente de ellas mismas. Podemos aceptar incluso, como Putnam pone de manifiesto en su ejemplo de un mundo carnapiano formado por tres objetos, que lo que consideramos individuos constituyentes del mundo real es algo que depende de nuestros esquemas conceptuales (en el ejemplo de Putnam el mundo puede estar formado por tres individuos o por siete u ocho, según el sistema lógico que adoptemos). Pero de ahí no tenemos por qué inferir que carezca de sentido suponer que algunas de las infi-

nitas representaciones del mundo que podemos construir pueda representar *mejor* la estructura objetiva del mundo real que otras muchas (también infinitas) representaciones lógicamente posibles.

En realidad lo único que realmente diferencia al realismo tentativo que proponemos del realismo interno que propone Putnam es que el primero asume la hipótesis metafísica que comparten tanto el realismo del sentido común como el realismo científico, sin tener que asumir al mismo tiempo la tesis ingenua de la verdad como correspondencia (o copia) del mundo real ni mucho menos la tesis metafísica de la existencia de una representación completa de la realidad (la tesis del Ojo de Dios), mientras que, si seguimos las argumentaciones de Putnam, parecería que ambos tipos de afirmaciones deberían ser siempre aceptadas o rechazadas en bloque.

Desde luego este «realismo tentativo» no está exento de problemas. Debería aclararse, por ejemplo, qué quiere decir exactamente una «representación adecuada pero parcial e incompleta» de la realidad, frente a una concepción estándar o ingenua de la verdad como correspondencia con los hechos. Pero al menos tiene esta virtud: descubrir que el concepto de verdad parcial e incompleta puede tener una importancia fundamental para nuestra imagen filosófica del mundo y de nuestro lugar en el mundo, importancia que una teoría como la del realismo interno de Putnam podría llevarnos a menospreciar injustamente.

Volvamos al comienzo de estos comentarios. Los discursos de los filósofos están llenos de términos abstrusos y de argumentaciones complicadas, hasta el punto de que con frecuencia parece que todo lo que dicen se refiere solamente a asuntos que sólo

a ellos interesan. Es una de las características de la filosofía académica, o escolástica, como Kant la llamaba. Pero fue el propio Kant quien estableció la distinción entre la filosofía académica y la filosofía mundana o universal, es decir entre la filosofía tal como la practican los filósofos y la filosofía que interesa a todo el mundo en tanto que persona racional. Lo difícil es tender un puente entre las dos: ser capaz de usar los conceptos y los argumentos técnicos del filósofo profesional de tal manera que sirvan para iluminar los problemas filosóficos que de forma implícita o explícita interesan al hombre de la calle. Y a la inversa, ser capaz de introducir en la discusión técnica de los filósofos cuestiones relevantes para la vida humana. Como muy bien señala James Conant, compilador del último libro de Putnam (*Realism with a Human Face*), ésta es precisamente una de las características del trabajo filosófico de nuestro autor, y es especialmente apreciable en el libro cuya versión castellana ahora presentamos. Espero que también estos torpes comentarios iniciales hayan contribuido a despertar en el lector, ante la obra que tiene en sus manos, la misma actitud rigurosa, crítica, perspicaz y creativa de la que su autor hace gala. Al fin y al cabo, en la vida cotidiana –y menos aún en el laboratorio científico– rara vez utilizamos la idea de verdad absoluta y completa; por el contrario nos basta con un concepto impreciso de verdad aproximada, tentativa y parcial. Salvo quizá en aquellas ocasiones en las que ejercemos de filósofos y empezamos a argumentar sobre valores absolutos, verdades eternas y distinciones incontrovertibles entre lo que es, lo que puede ser y lo que debe ser verdadero. Pero es posible que en esos casos estemos en realidad incorporando a

nuestro mundano discurso nuestra propia versión, imprecisa y poco rigurosa, del arsenal de conceptos técnicos que hemos ido heredando de los filósofos profesionales, entre otras razones porque la mayoría de éstos han prestado poca atención a los conceptos que realmente necesitamos.

***

Coda: Creo que debemos agradecer a Margarita Vázquez y a Manuel Liz el cuidado que han puesto en su traducción de la obra de Putnam al castellano. En mi caso además quiero agradecerles también sus comentarios y sugerencias a la versión inicial de esta introducción, algunos de los cuales he tenido en cuenta en la redacción definitiva. Otros quedan por el momento a la espera de que podamos seguir la discusión en alguna ocasión vitalmente más propicia (por lo que a mi respecta) para la plácida, abstrusa y apasionante especulación académica.

MIGUEL A. QUINTANILLA
Universidad de Salamanca

LAS MIL CARAS DEL REALISMO

PREFACIO

Cuando escribí *Reason, Truth and History*, describí mi propósito como el de romper el collar de fuerza que cierto número de dicotomías ponen sobre nuestro pensamiento, destacando entre ellas la dicotomía entre las perspectivas «objetivas» y «subjetivas» de la verdad y la razón. Describí mi concepción así (pág. xi): «Presentaré una perspectiva en la cual la mente no "copia" simplemente un mundo que puede ser descrito por "una teoría verdadera". Sin embargo, mi perspectiva no es una perspectiva en la cual la mente constituya el mundo (o lo componga sometida a constricciones impuestas por "cánones metodológicos" y "datos de los sentidos" independientes de la mente). Si tuviéramos que utilizar un lenguaje metafórico, entonces la metáfora sería ésta: la mente y el mundo, en común, constituyen la mente y el mundo».

La invitación a dar las Conferencias Paul Carus* en la reunión de diciembre de 1985 de la Asociación Filosófica Americana** en Washington, D.C., me ofreció la oportunidad de especificar con más profundidad la alternativa que veo, por un lado, a las perspectivas realistas metafísicas de la realidad y de la verdad y, por otro, a las del relativismo cultural. En el libro anterior, describí las concepciones habi-

* Paul Carus Lectures. [T.]

** American Philosophical Association, APA. [T.]

tuales de la verdad como perspectivas «alienadas» que hacen que se pierda una parte u otra de uno mismo y del mundo; en estas conferencias, he intentado desarrollar este comentario y explicar la conexión existente entre una concepción no alienada de la verdad y una concepción no alienada del florecimiento humano.

Es costumbre que las Conferencias Carus se publiquen en forma ampliada –en ocasiones, varias veces la longitud de las conferencias que fueron realmente impartidas–. Aquí, he intentado permanecer fiel a las conferencias que di en Washington con un cambio significativo: he insertado una conferencia (la actual Conferencia II) que no fue realmente leída en Washington. Me ha parecido preferible mantener el formato de conferencia a una reescritura que eliminaría la sensación real de una persona hablando a otras personas cara a cara.

Estoy en deuda con Daniel Álvarez, Thomas Carlson, Marcus Singer y Kenneth Winkler por sus recomendaciones históricas y por señalar ciertos errores. Como siempre, estoy en deuda, de un modo u otro, con todos mis colegas del Departamento de Filosofía de Harvard. Dieter Henrich (de Harvard y Munich), cuyas conferencias sobre Kant en Harvard han sido una inmensa fuente de inspiración, merece una atención especial. Si los nombres de Goodman, Quine y Rawls aparecen en estas páginas, es sólo una pequeña muestra de mi deuda con ellos. Agradezco particularmente a Burton Dreben muchas sugerencias que, según creo, han mejorado este trabajo. Y, una vez más, tengo que agradecer a Ruth Anna Putnam mucho más que su estímulo y consejo, pero ciertamente también esto.

CONFERENCIA I

¿QUEDA TODAVÍA ALGO POR DECIR ACERCA DE LA REALIDAD Y LA VERDAD?

El hombre de la calle, nos recuerda Eddington, ve una mesa como algo «sólido» –es decir, *principalmente* como materia sólida–. Sin embargo, la física ha descubierto que la mesa es, en su mayor parte, espacio vacío: que la distancia entre las partículas es inmensa en relación con el radio del electrón o con el núcleo de uno de los átomos que componen la mesa. Una reacción ante este estado de cosas, la reacción de Wilfrid Sellars,[1] consiste en negar rotundamente que haya mesas tal como normalmente las concebimos (aunque él escoge como ejemplo un cubito de hielo en vez de una mesa). La concepción que tiene el sentido común acerca de objetos materiales corrientes de tamaño medio, tales como mesas y cubitos de hielo (la «imagen manifiesta»), es simplemente *falsa* desde la perspectiva de Sellars (aunque tenga, al menos, algún valor cognitivo –existen objetos reales que las «mesas» y los «cubitos de hielo» de la imagen manifiesta «representan», según Sellars, aunque esos objetos reales no sean las mesas y los cubitos de hielo del hombre corriente–). No estoy de acuerdo con este enfoque de Sellars, pero espero que él me perdonará si utilizo éste, así como el fenómeno de su aparición en la escena filosófica, para ilustrar ciertas características del debate filosófico sobre el «realismo».

1. *Science, Perception, and Reality*, Atlantic Highlands, NJ: Humanities Press, 1963.

En primer lugar, esta perspectiva ilustra el hecho de que el Realismo con una «R» mayúscula no siempre da lo que el ingenuo espera de él. Si hay algún atractivo del Realismo que sea enteramente legítimo es la apelación al sentimiento del sentido común de que *por supuesto* hay mesas y sillas, y cualquier filosofía que nos diga que realmente no existen –que realmente sólo hay datos de los sentidos, o sólo «textos», o lo que sea–, es algo más que disparatada. El Realismo me recuerda al seductor en el melodrama pasado de moda. En los melodramas de la década de 1890, el seductor siempre prometía a la inocente doncella cosas diversas que nunca cumplía cuando llegaba el momento. En este caso, el realista (el malvado seductor) promete al sentido común (la inocente doncella) que la rescatará de sus enemigos (idealistas, kantianos y neokantianos, pragmatistas y el temible, autodenominado «irrealista» Nelson Goodman) quienes (el realista dice) quieren privarla de sus queridos cubitos de hielo y sillas. Enfrentada a este terrible porvenir, la bella doncella elige, por supuesto, la compañía del realista del sentido común. Pero, cuando han viajado juntos un rato, el «realista científico» anuncia que lo que la doncella va a conseguir *no son* sus cubitos de hielo, sus mesas y sillas. De hecho, todo lo que *realmente* hay –le dice el realista científico durante el desayuno– es lo que la «ciencia acabada» dirá que hay –sea lo que sea–. Ella es abandonada con un pagaré para «ella no sabe qué», y con la garantía de que aunque *no haya* mesas y sillas, hay aún algunas *Dinge an sich* que su «imagen manifiesta» (o su «física natural», en expresión de algunos realistas científicos) «representa». Algunos dirán que la dama ha sido engañada.

Así, está claro que el rótulo «Realismo» puede ser reclamado por –o ser propenso a– al menos dos actitudes filosóficas muy distintas (y, de hecho, a muchas más). El filósofo que pretende que sólo «existen realmente» los objetos científicos y que gran parte del mundo del sentido común, si no todo, es simple «proyección» reclama ser un realista, pero también lo hace el filósofo que insiste en que *realmente hay* sillas y cubitos de hielo (y en que algunos de estos cubitos de hielo realmente son *de color rosa*), y estas dos actitudes, estas dos imágenes del mundo pueden, conducir, y han conducido, a muy variados programas en filosofía.

Husserl[2] sigue la pista de la primera línea de pensamiento, la línea que niega que «realmente hay» objetos del sentido común, hasta Galileo, y con mucha razón. La actual concepción occidental del mundo depende, de acuerdo con Husserl, de una nueva manera de concebir los «objetos externos», la de la física matemática. Un objeto externo se concibe como una congregación de partículas (por los atomistas) o como algún tipo de distorsión ampliada (en el siglo XVII, un «vórtice» y, más tarde, una colección de «campos»). De cualquier forma, la mesa frente a mí (o el objeto que yo me «represento como» una mesa) se describe, como dice Husserl, mediante «fórmulas matemáticas». Y esto, señala, es lo que trajo sobre todo la revolución galileana al pensamiento occidental: la idea del «mundo externo» como algo cuya descripción verdadera, cuya descripción «tal como es en sí mismo», consiste en fórmulas matemáticas.

2. *The Crisis of the European Sciences and Transcendental Phenomenology*, traducido por David Carr, Evanston: Northwestern University Press, 1970.

Es importante, para esta forma de pensar, que ciertas propiedades familiares de la mesa –su tamaño, forma y lugar– sean propiedades «reales», descriptibles, por ejemplo, en el lenguaje de la geometría analítica de Descartes. Otras propiedades, sin embargo, las así llamadas propiedades «secundarias», de las cuales el *color* es un ejemplo clave, *no* son tratadas como propiedades reales en el mismo sentido. De ninguna propiedad «actualmente ejemplificada» (no disposicional) por ese enjambre de moléculas (o por esa región espacio-temporal), reconocida en la física matemática, puede decirse que sea lo que todos nosotros llamábamos su *color*.

¿Qué pasa con las propiedades disposicionales? A menudo se sostiene que el color es simplemente una función de *la reflexión de la luz*, esto es, de la disposición de un objeto (o de la superficie de un objeto) para absorber selectivamente ciertas longitudes de onda de luz incidente y reflejar otras. Pero, verdaderamente, esto no tiene mucho que ver con la realidad de los colores. Las investigaciones recientes no sólo han mostrado que esta consideración es demasiado simple (porque cambios de reflexión en los bordes pasan a jugar un importante papel en la determinación de los colores que vemos), sino que la reflexión, en sí misma, no tiene una explicación física uniforme. Una estrella roja, una manzana roja y un vaso lleno de agua coloreada de rojo son rojos por razones físicas bastante distintas. De hecho, puede haber un número infinito de condiciones físicas diferentes que podrían producir la disposición de reflejar (o emitir) luz roja y absorber luz de otras longitudes de onda. Una propiedad disposicional cuya «explicación» no disposicional subyacente sea tan poco uniforme simple-

mente no es susceptible de ser representada como una función matemática de las variables dinámicas. Y éstas –las variables dinámicas– son los parámetros que esta forma de pensar considera como las «características» de los «objetos» externos.

Otro problema[3] radica en que los *matices* resultan ser algo mucho más subjetivo de lo que pensábamos. De hecho, cualquier tono en el cuadro de colores correspondiente a la parte verde del espectro será clasificado como «verde estándar» por algún sujeto, aunque esté al final del extremo «amarillo-verde» o del extremo «azul-verde».

Resumiendo, ninguna «característica» reconocida por esta forma de pensar –ninguna «función de las variables dinámicas con las propiedades adecuadas»– se corresponde con una propiedad tan familiar de los objetos como *rojo* o *verde*. La idea de que existe una propiedad que todos los objetos rojos tienen en común –la misma en todos los casos–, y otra propiedad que todos los objetos verdes tienen en común –la misma en todos los casos–, es una especie de ilusión en la perspectiva que hemos ido aceptando cada vez más desde la época de Descartes y Locke.

Sin embargo, Locke y Descartes nos dieron un sofisticado sustituto para nuestra noción precientífica del color; un sustituto que, quizás, a la mayoría de la gente le ha llegado a parecer simple «sentido común post-científico». Este sustituto involucra la idea de un dato de los sentidos (con la salvedad de que, en el vocabulario de los siglos XVII y XVIII, los datos de los sentidos eran llamados «ideas» o «im-

3. Véase, de C. L. Hardin, «Are "Scientific" Objects Colored?», en *Mind*, XCIII, n. 22 (octubre, 1964), 491-500.

presiones»). El jersey rojo que yo veo no es rojo a la manera como pensaba que lo era (no existe ninguna «magnitud física» que sea su rojez), pero tiene una disposición (un poder, en el lenguaje de los siglos XVII y XVIII) para afectarme de una cierta forma –para hacerme tener datos de los sentidos–. Y éstos, los datos de los sentidos, verdaderamente tienen una clase de «rojez» simple, uniforme y no disposicional.

Ésta es la famosa imagen, la imagen dualista del mundo físico y sus cualidades primarias, por un lado, y de la mente y sus datos de los sentidos, por otro, a causa de la cual los filósofos se han estado peleando, como dice Husserl, desde los tiempos de Galileo. Y Husserl tiene la intuición –como la tuvo William James, quien influyó en Husserl– de que esta imagen es desastrosa.

Pero, ¿por qué debemos verla como desastrosa? Seguramente alguna vez fue chocante, pero, como ya he dicho, en la actualidad es ampliamente aceptado como «sentido común post-científico». ¿Qué es lo que está *realmente* equivocado en esta imagen?

Por alguna razón, la *solidez* está prácticamente en el mismo barco que el color. Si los objetos no tienen color tal como «inocentemente» parecen tenerlo, tampoco tienen solidez tal como «inocentemente» parecen tenerla.[4] Es esto lo que conduce a Sellars a decir que objetos del sentido común, del tipo de los cubitos de hielo, no existen realmente. ¿Cuál *es* nuestra concepción de un típico objeto del senti-

4. La noción de sentido común de «solidez» no debería ser confundida con la noción de los físicos de estar «en estado sólido». Por ejemplo, una duna de arena está en «estado sólido», pero no es sólida en el sentido ordinario del término; mientras que una botella de leche puede ser sólida, aunque la mayor parte de su contenido no esté en estado sólido.

do común aparte de algo sólido (o líquido) que exhibe ciertos colores? Lo que realmente hay, en la metafísica científica de Sellars, son objetos de la física matemática, por un lado, y «sensaciones desnudas», por otro. Ésta es justamente la imagen que acabo de describir como «desastrosa»; la imagen que niega precisamente el tipo de realismo del hombre común, su realismo sobre las mesas y las sillas.

La réplica (la réplica que hará un filósofo que acepta la imagen post-galileana) me parece obvia: «Te sientes simplemente nostálgico de un mundo más antiguo y más simple. Esta imagen funciona; nuestra aceptación de ella es una "inferencia basada en la mejor explicación". No podemos considerar como objeción una perspectiva que no conserve todo lo que el hombre de la calle alguna vez creyó falsamente».

Aunque se trate de una inferencia basada en la mejor explicación, es bastante extraña. ¿Cómo discurre la explicación familiar de lo que ocurre cuando «veo algo rojo»? La luz choca contra el objeto (digamos, un jersey) y se refleja hacia mi ojo. Hay una imagen en la retina (Berkeley conocía algunas cosas acerca de las imágenes en la retina, como también Descartes, aunque el carácter ondulatorio de la luz no se entendió bien hasta mucho más tarde). Se producen ciertos impulsos nerviosos (Descartes sabía que había alguna clase de transmisión a través de los nervios, aunque estaba equivocado acerca de su naturaleza –y tampoco resulta claro que nosotros conozcamos su naturaleza, ya que de nuevo se discute la relevancia de las transmisiones químicas de neurona a neurona, en contraste con las eléctricas–). Ocurren cosas en el cerebro, algunas de las cuales entendemos gracias al trabajo de Hubel y Wiesel, David Marr y otros. Y, en-

tonces –ésta es la parte misteriosa–, de alguna manera surge un «dato de los sentidos» o una «sensación desnuda». ¿Es *esto* una *explicación*?

Una «explicación» que involucra conexiones de un tipo que no entendemos en absoluto (conexiones en «suspenso nomológico» las llamó Herbert Feigl[5]), y respecto de las cuales no tenemos ni siquiera el esbozo de una teoría, es una explicación por medios algo más oscuros que el fenómeno que ha de ser explicado. Como ha sido señalado por pensadores tan diferentes uno de otro como William James, Husserl y John Austin, cada uno de los apartados del relato acerca del dato de los sentidos es una suposición –teoría–; y una teoría de una clase de lo más peculiar. Y, sin embargo, el papel epistemológico que la filosofía tradicional supone que juegan los «datos de los sentidos» exige de ellos que sean lo que es «dado» que sean, *aquello de lo cual estemos absolutamente seguros independientemente de la teoría científica*. El tipo de realismo científico que hemos heredado del siglo XVII aún no ha perdido todavía todo su prestigio, pero nos ha hecho cargar con una imagen desastrosa del mundo. Ya va siendo hora de buscar una imagen diferente.

Propiedades intrínsecas: las disposiciones

Quiero sugerir que el problema con la imagen «objetivista» del mundo (por usar el término de

5. «The "Mental" and the "Physical"», en *Minnesota Studies in the Philosophy of Science*, vol. II, *Concepts, Theories and the Mind-Body Problem*, por Feigl, Scriven y Maxwell, Minneapolis: University of Minnesota Press, 1958, 370-497.

Husserl para esta clase de realismo científico) es mucho más profundo que la postulación de «datos de los sentidos»; los datos de los sentidos son, por decirlo así, los síntomas visibles de una enfermedad del sistema, como las marcas de pústulas en el caso de la viruela. La raíz profunda de la enfermedad en el sistema descansa en la noción de una propiedad «intrínseca», una propiedad que algo tiene «en sí mismo», al margen de cualquier contribución hecha por el lenguaje o la mente.

Esta noción, y la noción correlativa de una propiedad que es meramente «apariencia», o meramente algo que nosotros «proyectamos» en el objeto, ha resultado ser tremendamente robusta, a juzgar por el atractivo que ha tenido para diferentes tipos de filósofos. A pesar de sus profundos desacuerdos, todas las corrientes de la filosofía que aceptaron el grupo de problemas del siglo XVII –idealistas subjetivos así como dualistas y materialistas– aceptaron la distinción, incluso estando en desacuerdo sobre su aplicación. Un idealista subjetivo diría que sólo hay datos de los sentidos (o mentes y datos de los sentidos, en algunas versiones), y que «rojo» es una propiedad intrínseca de estos objetos, mientras que la persistencia (estar allí incluso cuando no miramos) es algo que nosotros «proyectamos»; un dualista o un materialista diría que los objetos «externos» tienen persistencia como una propiedad intrínseca, pero que rojo es, en su caso, algo que nosotros «proyectamos». No obstante, todos estos filósofos *tienen* la distinción. Incluso Kant, que expresó serias dudas sobre ella en la primera Crítica (hasta el punto de decir que la noción de un *Ding an sich puede* ser «vacía»), hace un abundante uso de ella en la segunda Crítica.

Dejando de lado la concepción berkeleyana (que realmente no hay en absoluto objetos externos) como una forma aberrante de la perspectiva del siglo XVII, podemos decir que todos los filósofos restantes aceptan la explicación de la «rojez» y la «solidez» que he estado describiendo; éstas no son «propiedades intrínsecas» de las cosas externas a las cuales se las adscribimos, sino más bien (en el caso de las cosas externas) disposiciones para afectarnos de ciertas maneras –para producir ciertos datos de los sentidos en nosotros o, como dirían los filósofos materialistas, para producir ciertas clases de «estados» en nuestros cerebros y sistemas nerviosos–. La idea de que estas propiedades están «en» las cosas mismas, como propiedades intrínsecas, es una «proyección» espontánea.

El talón de Aquiles de esta historia es la noción de disposición. Para señalar los problemas que surgen –y que han preocupado a muchas mentes filosóficas de primera línea, empezando por la de Charles Peirce–, permítanme introducir un término técnico (no introduciré mucha terminología en esta conferencia, ¡lo prometo!). Llamaré *disposición estricta* a una disposición que algo tiene para hacer algo, *no importa lo que ocurra*. Llamaré *disposición ceteris paribus* a una disposición para hacer algo bajo «condiciones normales». Quizás sería prudente dar ejemplos.

La disposición de los cuerpos con una masa en reposo no igual a cero de viajar a velocidades por debajo de la velocidad de la luz es una disposición *estricta*; es físicamente imposible para un cuerpo con una masa en reposo no igual a cero viajar a la velocidad de la luz. Por supuesto, la noción de «disposición estricta» presupone la noción de «necesi-

dad física», como este ejemplo ilustra, pero es una noción que le estoy permitiendo al «realista científico», al menos en consideración al argumento. ¿Qué ocurre con la disposición del azúcar para disolverse en agua?

Ésta no es una disposición estricta, ya que el azúcar que se coloca en agua ya saturada con azúcar (o, incluso, con otros productos químicos apropiados) no se disolverá. ¿Es, entonces, una disposición estricta la disposición del azúcar para disolverse en *agua químicamente pura*?

He aquí otra disposición no estricta: el primer contraejemplo que mencionaré proviene de la termodinámica. Supongamos que arrojo un terrón de azúcar en agua y que el terrón de azúcar se disuelve. Consideremos azúcar que está en agua, pero de tal forma que, aunque la situación sea idéntica a la situación que acabo de imaginar (el azúcar está disuelto en el agua) con respecto a la posición de cada partícula, y también con respecto al valor numérico del momento de cada partícula, todos los vectores de momento tienen una orientación exactamente opuesta respecto de los que ahora tenemos. Se trata de un ejemplo famoso: lo que ocurre en el ejemplo es que el azúcar, en lugar de permanecer disuelto, simplemente forma un terrón de azúcar que, espontáneamente, ¡salta fuera del agua! Puesto que a todo estado normal (todo estado en el cual el azúcar se disuelve) le corresponde un estado en el cual «se des-disuelve», vemos que hay infinitamente muchas condiciones físicamente posibles en las cuales el azúcar «se des-disuelve» en vez de permanecer disuelto. Por supuesto, todos son estados en los que la entropía disminuye; pero eso no es imposible, ¡sólo extremadamente improbable!

¿Diremos, entonces, que el azúcar tiene una disposición estricta a disolverse a menos que se encuentre en una condición en la cual tenga lugar una disminución de la entropía? No, porque si el azúcar se coloca en agua y se produce una congelación repentina, el azúcar no se disolverá si la congelación tiene lugar lo suficientemente rápido...

De hecho, lo que podemos decir es que bajo condiciones *normales* el azúcar se disolverá si es colocado en agua. Y no hay ninguna razón para pensar que todas las diferentes condiciones no normales (incluyendo extraños estados cuánticos, extrañas fluctuaciones locales en el espacio-tiempo, etc.), bajo las cuales el azúcar no se disolvería si fuera colocado en agua, puedan recogerse en una fórmula cerrada del lenguaje de la física fundamental.

¡Éste es exactamente el problema que hemos observado antes en conexión con la rojez y la solidez! Si son las propiedades «intrínsecas» de las cosas «externas» las que podemos representar mediante fórmulas en el lenguaje de la física fundamental, por medio de «adecuadas funciones de las variables dinámicas», entonces la *solubilidad* tampoco es una propiedad «intrínseca» de ninguna cosa externa. Y, de manera similar, tampoco lo es ninguna disposición «en igualdad de circunstancias». Los poderes, por usar el lenguaje del siglo XVII, tienen que oponerse a las propiedades que las cosas tienen «en sí mismas», y ser cuidadosamente distinguidos de ellas.

Propiedades intrínsecas: la intencionalidad

Bien, ¿qué hay de malo en ello? ¿Por qué no deberíamos decir que las disposiciones (o, al menos,

las disposiciones «en igualdad de circunstancias», tales como la solubilidad) tampoco están «en las cosas mismas» sino, más bien, en algo que nosotros «proyectamos» en esas cosas? Los filósofos que hablan de esta forma raramente se paran a decir, si es que lo hacen alguna vez, qué se supone que es la *proyección* misma. ¿De dónde proviene, en este esquema, la habilidad de la mente para «proyectar» algo en algo?

La proyección consiste en pensar que algo tiene propiedades que no tiene, pero que podemos imaginar (quizás porque alguna cosa con la que estamos familiarizados realmente las tiene), sin ser conscientes de que esto es lo que estamos haciendo. Es, pues, una especie de *pensamiento* –pensamiento sobre algo–. ¿Tiene algo que decirnos acerca del pensamiento (o, como los filósofos dicen, acerca de la «intencionalidad», es decir, acerca del *tratar sobre algo*) la conocida imagen «objetivista»?

Descartes, efectivamente, pretendió que debería hacerlo. Su tesis consistía en que hay dos sustancias fundamentales –la mente y la materia–, no una sola, y en que, de modo equivalente, debería haber dos ciencias fundamentales: la física y la psicología. Pero hemos dejado por completo de pensar en la mente como en una «sustancia» separada. Y, contrariamente a las esperanzas de Descartes, nunca ha llegado a surgir una psicología como «ciencia fundamental» que explique la naturaleza del pensamiento (incluyendo cómo los pensamientos pueden ser verdaderos o falsos, justificados o no justificados, sobre algo o sobre nada). En consecuencia, explicar las características del mundo del sentido común, incluyendo el color, la solidez, la causalidad (incluyo la causalidad porque, si las disposiciones

son «proyecciones», la noción del sentido común de «la causa» de algo es una «proyección»; depende de la noción de «condiciones normales» exactamente en el mismo sentido), en términos de una operación mental llamada «proyección» es justamente explicar cada una de estas características del mundo del sentido común en función del *pensamiento*.

Pero, ¿no era esto lo que se les acusaba de hacer a los idealistas? Ésta es la paradoja que señalé al principio de esta conferencia. Por lo que se refiere al mundo del sentido común (el mundo en el que tenemos experiencia de nosotros mismos como *viviendo* en él, que es por lo que Husserl lo llamó *Lebenswelt*), el efecto de lo que se llama «realismo» en filosofía consiste en negar su realidad objetiva, en convertirlo simplemente en *pensamiento*. Son los filósofos que, en un sentido u otro, se mantienen en la tradición neokantiana –James, Husserl, Wittgenstein– los que mantienen que las mesas y sillas del sentido común, las sensaciones y los electrones, son *igualmente reales*, y no los realistas metafísicos.

Hoy en día, algunos realistas metafísicos dirían que no necesitamos una perfecta ciencia de la psicología para dar cuenta del pensamiento y de la intencionalidad, ya que el problema se resuelve mediante alguna teoría filosófica: mientras tanto, otros pretenden que una «ciencia cognitiva» perfecta, basada en el «modelo del ordenador», resolverá el problema por nosotros en un futuro más o menos próximo o distante. Obviamente, hoy no tengo tiempo para examinar en detalle estas sugerencias, pero indicaré brevemente por qué creo que ninguna de ellas resiste un examen profundo.

¿Por qué la intencionalidad es un problema tan insoluble?

El problema, por decirlo brevemente, es que el pensamiento mismo ha sido tratado por la filosofía, cada vez en mayor medida, como una «proyección» que remite su pedigrí al siglo XVII. La razón es clara: no hemos tenido éxito a la hora de dar algún contenido a la teoría de que el pensamiento es, justamente, una propiedad primitiva de una misteriosa «sustancia»: la mente. Como Kant señaló en la primera Crítica, no tenemos ninguna teoría de esta sustancia o de sus poderes, y ninguna perspectiva de tener una. Si, a diferencia del Kant de la primera Crítica (tal como yo leo la *Crítica de la razón pura*), insistimos en ser fieles a los supuestos «objetivistas» fundamentales, el único camino que entonces podemos tomar es el de que *los fenómenos mentales tienen que ser, en algún sentido, estrechamente derivados de fenómenos físicos*, como Diderot y Hobbes ya habían propuesto. Por «supuestos objetivistas fundamentales» entiendo a) el supuesto de que se puede trazar una clara distinción entre las propiedades que las cosas tienen «en sí mismas» y las propiedades que «nosotros proyectamos», y b) el supuesto de que la ciencia fundamental –en singular, ya que sólo la física tiene hoy ese *status*– nos dice qué propiedades tienen las cosas «en sí mismas». (Aunque quisiéramos suponer, junto con Wilfrid Sellars, que las «sensaciones desnudas» –las cualidades sensibles fundamentales de la experiencia– no van a ser reducidas a física, sino que de alguna forma han de añadirse a la ciencia fundamental en algún siglo futuro, ello no afectaría mucho a la situación; Sellars no adelanta que la *intencionalidad* se

convertirá en algo que tenemos que añadir, en ese mismo sentido, a la física, sino más bien supone que una teoría del «uso de las palabras» es todo lo que se necesita para dar cuenta de ella.)

El objetivismo moderno se ha convertido simplemente en materialismo. Y el problema central para el materialismo consiste en «explicar la aparición de la mente». Pero, si «explicar la aparición de la mente» significa resolver el problema de Brentano, esto es, decir en términos *reductivos* qué *es* «pensar que hay muchos gatos en el vecindario» y qué *es* «recordar dónde está París», etc., ¿por qué deberíamos pensar ahora que *eso* es posible? Si se ha demostrado que reducir el color, la solidez o la solubilidad a la física fundamental es imposible, ¿por qué debería ser más fácil de resolver este programa reduccionista mucho más ambicioso?

Empezando a finales de la década de los cincuenta, yo mismo propuse un programa en filosofía de la mente que ha llegado a ser ampliamente conocido bajo el nombre «funcionalismo». Mi «funcionalismo» afirmaba que los seres pensantes son *composicionalmente plásticos*, esto es, que no hay ningún estado o evento físico (es decir, ninguna condición necesaria y suficiente expresable a través de una fórmula finita del lenguaje de la física fundamental de primer orden) para ser, ni siquiera, una ocurrencia *físicamente posible* (dejemos aparte la «posibilidad lógica» o la «posibilidad metafísica») de un pensamiento con un contenido proposicional dado, o un sentimiento de cólera, o un dolor, etc. *A fortiori*, las actitudes proposicionales, las emociones, los sentimientos, no son *idénticos a estados cerebrales*, *ni siquiera* a estados físicos caracterizados más ampliamente. Cuando propuse esta tesis, se-

ñalé que pensar sobre la mentalidad, la afectividad, etc., de un ser como aspectos de su *organización para realizar una función* nos permite reconocer que cualquier clase de «sistema» o de ser lógicamente posible podría ser consciente, exhibir mentalidad y afecto, etc., exactamente en el mismo sentido sin tener la misma materia (sin estar hecho, en absoluto, de materia en el sentido de las partículas elementales y de los campos electromagnéticos). Ya que seres con constituciones físicas muy diferentes (e incluso «no-físicas») podrían tener la misma organización funcional. Lo que queremos investigar es la naturaleza de la organización funcional humana (y animal), no la naturaleza de una misteriosa «sustancia» por un lado, o una simple información fisiológica adicional, por otro.

También propuse una teoría acerca de en qué consistía nuestra organización funcional, una teoría que actualmente he abandonado –era la teoría de que nuestra organización funcional era la de una máquina de Türing–. La he abandonado porque creo que hay buenos argumentos que muestran que los estados mentales no son sólo composicionalmente plásticos sino, también, *computacionalmente plásticos*. Lo que quiero decir con esto es que criaturas físicamente posibles que crean que hay muchos gatos en el vecindario, o lo que sea, pueden hacerlo mediante un *número indefinido de «programas» diferentes*. La hipótesis de que hay una condición necesaria y suficiente en términos computacionales (o en términos computacionales *cum* físicos) para la presencia de una creencia dada es poco realista, justamente en el mismo sentido de que la teoría de que hay una condición necesaria y suficiente, en términos fenomenalistas, para la presencia de una mesa

es poco realista. Tal condición tendría que ser infinitamente larga, y no construida de acuerdo con ninguna regla efectiva, ni siquiera conforme a una prescripción no efectiva que pudiéramos proponer sin usar los propios términos que han de reducirse. Ni siquiera creo que todos los *humanos* que tengan la misma creencia (en diferentes culturas, o con diferentes cuerpos de conocimiento y recursos conceptuales diferentes) tengan en común una característica física *cum* computacional que pueda ser «identificada con» esa creencia. El «nivel intencional», simplemente, no es reductible al «nivel computacional» más de lo que pueda serlo al «nivel físico».[6]

Si esto es correcto, entonces el objetivista tendrá que concluir que *también* la intencionalidad debe ser una mera «proyección». Pero, ¿cómo puede ningún filósofo pensar que esta sugerencia parezca siquiera tener sentido? Como vimos, ¡la propia noción de «proyección» *presupone* la intencionalidad!

Por extraño que parezca, y a pesar de su carácter absurdo, la idea de que el pensamiento *es* una mera proyección está siendo defendida por varios filósofos en los Estados Unidos y en Inglaterra. La fuerza de la tradición «objetivista» es tan intensa que algunos filósofos abandonarán las más profundas intuiciones que tenemos sobre nosotros-mismos-en-el-mundo antes que preguntarse (como hicieron Husserl y Wittgenstein) si la imagen completa no es un error. Por ello, en las últimas décadas del siglo XX nos encontramos con filósofos inteligentes[7] que pro-

6. Esto se argumenta en mi *Representation and Reality*, Montgomery, Vt: Bradford Books, próxima publicación.

7. D. C. Dennett, *Content and Consciousness*, Atlantic Highlands, NJ: Humanities Press, 1969.

ponen que la intencionalidad misma es algo que nosotros proyectamos adoptando una «perspectiva» hacia ciertas partes del mundo (¡como si «adoptar una perspectiva» no fuera, a su vez, una noción intencional!), filósofos inteligentes que pretenden que nadie tiene realmente actitudes proposicionales (creencias y deseos), que «creencia» y «deseo» son únicamente nociones de una falsa teoría llamada «psicología natural», y filósofos inteligentes que defienden que no existe una propiedad como la «verdad» ni una relación como la referencia, que «es verdad» no es más que una locución que usamos para «elevar el nivel del lenguaje». Uno de ellos –Richard Rorty, un pensador de gran profundidad– observa que está comprometido a rechazar las intuiciones que subyacen a cualquier clase de realismo[8] (y no sólo al realismo metafísico), aunque la mayoría de estos pensadores escriben como si estuvieran *salvando* el realismo (en su versión materialista), ¡abandonando la intencionalidad! ¡Es como si no hubiera ningún problema en decir «yo no niego que hay un mundo externo; lo que niego es que *pensemos* sobre él»! Créanlo, ésta es también la forma en que escribía Foucault. ¡La línea entre el relativismo *à la française* y la filosofía analítica parece ser más tenue de lo que los filósofos angloparlantes piensan! Algo bastante divertido es que la sobrecubierta de uno de los últimos ataques a la «psicología natural»[9] lleva un anuncio entusiasta en el cual un crítico explica la importancia del libro, que está dentro

8. *Philosophy and the Mirror of Nature*, Princeton: Princeton University Press, 1979.

9. Stephen Stich, *From Folk Psychology to Cognitive Science: The case Against Belief*, Cambridge, MA: MIT Press, 1983.

de la sobrecubierta, diciendo que ¡la mayoría de la gente *cree* que hay cosas tales como las creencias!

«El rastro de la serpiente humana está por todas partes»

Si el objetivismo del siglo XVII ha conducido a la filosofía del siglo XX a un callejón sin salida, la solución no es ni caer en un relativismo extremo, como la filosofía francesa ha estado haciendo, ni negar nuestro realismo de sentido común. *Hay* mesas, sillas y cubitos de hielo. Hay también electrones, regiones espacio-temporales, números primos, gente que es una amenaza para la paz del mundo, momentos de belleza y trascendencia y muchas otras cosas. Mi historia pasada de moda del seductor y la inocente doncella significaba un doble aviso; un aviso contra el abandono del realismo del sentido común y, simultáneamente, un aviso contra el supuesto de que la manera de hablar en el siglo XVII de «mundo externo» e «impresiones de los sentidos», «propiedades intrínsecas» y «proyecciones», etc., entrañaba, de alguna forma, un rescate de nuestro realismo de sentido común. El Realismo con una «R» mayúscula es, da pena decirlo, el enemigo, no el defensor, del realismo con una «r» minúscula.

Si es difícil ver esto es porque la tarea de superar la imagen del mundo del siglo XVII no ha hecho más que empezar. Pregunté –como título de esta conferencia– si aún había algo que decir, algo realmente nuevo que decir, sobre la realidad y la verdad. Si «nuevo» significa «completamente sin precedentes», sospecho que la respuesta es «no». Pero, si permitimos que William James pueda haber tenido

algo «nuevo» que decir, algo nuevo para *nosotros*, no solamente nuevo para su propio tiempo o, al menos, que pueda haber tenido un programa para la filosofía que sea, en parte, el programa adecuado, aunque no haya sido desarrollado de modo adecuado (y tal vez nunca sea completamente «desarrollado»); si permitimos que Husserl, Wittgenstein y Austin puedan haber compartido algo del mismo programa, a pesar de que ellos también, cada uno a su manera, fracasaran en presentarlo adecuadamente, entonces aún hay algo nuevo, algo *inacabado e importante* que decir sobre la realidad y la verdad. Y esto es lo que yo creo.

La clave para desarrollar el programa de conservar el realismo de sentido común a la vez que se evitan los absurdos y antinomias del realismo metafísico en todas sus variedades (marca X: materialismo; marca Y: idealismo subjetivo; marca Z: dualismo...) es algo que yo he llamado *realismo interno*. (¡Debería haberlo llamado realismo pragmático!). El realismo interno es, en el fondo, únicamente la insistencia en que el realismo *no* es incompatible con la relatividad conceptual. Se puede ser *al mismo tiempo* un realista *y* un relativista conceptual. El realismo (con una «r» minúscula) ha sido ya introducido; como decíamos, es la perspectiva que toma nuestro esquema del sentido común familiar, tanto como nuestros esquemas científicos, artísticos y otros, en sentido literal, sin ayudarse de la noción de la cosa «en sí misma». Pero, ¿qué es la relatividad conceptual?

Relatividad conceptual suena como «relativismo», pero no tiene ninguna de las implicaciones del relativismo del tipo «no existe ninguna verdad que descubrir... "verdadero" no es más que un nombre para aquello en lo que un grupo de gente puede es-

tar de acuerdo». Un sencillo ejemplo ilustrará lo que quiero decir. Considérese «un mundo con tres individuos» (Carnap a menudo usaba ejemplos como éste cuando estábamos haciendo juntos lógica inductiva a principios de los cincuenta), *x1*, *x2*, *x3* ¿Cuántos *objetos* hay en ese mundo?

Bien, dije «considérese un mundo con sólo tres individuos», ¿no? Por lo tanto, ¿no tiene que haber allí tres objetos? ¿Puede haber entidades no abstractas que no sean «individuos»?

Una respuesta posible es «no». Podemos identificar «individuo», «objeto», «particular», etc., y no encontrar nada absurdo en un mundo con sólo tres individuos que sean independientes, «átomos lógicos» sin ninguna relación. Pero existen doctrinas lógicas perfectamente respetables que conducen a resultados diferentes.

Supóngase, por ejemplo, que yo, como algunos lógicos polacos, creo que para dos particulares cualesquiera hay un objeto que es su suma. (Ésta es la asunción básica de la «mereología», el cálculo de partes y todos inventado por Lezniewski.) Si ignoro, por un momento, el así llamado «objeto nulo», entonces encontraré que el mundo de «tres individuos» (como Carnap puede haberlo tenido, al menos cuando estaba haciendo lógica inductiva) realmente contiene *siete* objetos:

` Mundo 1	Mundo 2
x1, x2, x3	x1, x2, x3, x1 + x2, x1 + x3, x2 + x3, x1 + x2 + x3
(Un mundo al estilo de Carnap)	(El «mismo» mundo al estilo de los lógicos polacos)

Algunos lógicos polacos dirían también que hay un «objeto nulo» que ellos cuentan como parte de todo objeto. Si aceptamos esta sugerencia, y añadimos este individuo (llamémosle O), entonces diríamos que el mundo de Carnap contiene *ocho* objetos.

Ahora, la manera realista metafísica clásica de enfrentarse con tales problemas es bien conocida. Consiste en decir que hay un único mundo (piénsese en él como en un trozo de masa de pastelería) que podemos cortar en piezas de diferentes formas. Pero, esta metáfora del «molde de pastas» echa por tierra la pregunta ¿cuáles son las «partes» de esta «masa»? Si la respuesta es que O, x1, x2, x3, x1 + x2, x1 + x3, x2 + x3, x1 + x2 + x3 son todas las diferentes «piezas», entonces no tenemos una descripción *neutral*, sino más bien una descripción *partidista* –¡justamente la descripción del lógico de Varsovia!–. Y no es casualidad que el realismo metafísico no pueda realmente reconocer el fenómeno de la relatividad conceptual –ya que tal fenómeno se convierte en el hecho de que *los mismos conceptos lógicos primitivos, y en particular las nociones de objeto y existencia, tienen una multitud de diferentes usos más que un «significado» absoluto*.

Un ejemplo que es históricamente importante, aunque más complejo que el que acabamos de dar, es la antigua disputa acerca del *status* ontológico del plano euclidiano. Imagínense un plano euclidiano. Piénsese en los puntos del plano. ¿Son éstos *partes* del plano, como Leibniz pensó?, o ¿son «meros límites», como dijo Kant?

Si se dice, en *este* caso, que se trata de «dos maneras de cortar la misma masa», entonces se debe

admitir que lo que es una *parte* del espacio, en una versión de los hechos, es una entidad abstracta (digamos, un conjunto de esferas convergentes –aunque no hay, naturalmente, una forma única de construir puntos como límites–) en la otra versión. Pero, entonces, se tendrá que conceder que, al menos, la distinción entre qué entidades son «entidades abstractas» y qué entidades son «objetos concretos» sea relativa a una versión. Los realistas metafísicos continúan, hasta el momento presente, discutiendo si los puntos (puntos espacio-temporales, en la actualidad, más que puntos en un plano o en un espacio tridimensional) son individuos o propiedades, particulares o simples límites, etc. Mi opinión es que Dios mismo, si consintiera en responder a la pregunta «¿existen realmente los puntos o son meros límites?», diría «no lo sé»; no porque Su omnisciencia sea limitada, sino porque hay un límite más allá del cual las preguntas no tienen sentido.

Un último punto antes de que deje estos ejemplos: *dada* una versión, la pregunta «¿cuántos objetos hay?» tiene una respuesta; ésta es «tres» en el caso de la primera versión («el mundo de Carnap») y «siete» (u «ocho») en el caso de la segunda versión («el mundo del lógico polaco»). Una vez que aclaramos cómo estamos usando «objeto» (o «existe»), la pregunta (¿cuántos objetos existen?) tiene una respuesta que no es, en absoluto, un asunto de «convención». Por eso digo que este tipo de ejemplo no da apoyo al relativismo cultural *radical*. Nuestros conceptos pueden ser relativos a una cultura, pero de aquí no se sigue que la verdad o la falsedad de cualquier cosa que digamos usando esos conceptos sea simplemente «decidida» por la cultura.

Aunque la idea de que hay un punto arquimediano, o un uso de «existe» inherente al mundo mismo, desde el cual la pregunta «¿cuántos objetos existen *realmente*?» cobra sentido, es una ilusión.

Si esto es correcto, entonces ha de ser posible ver cómo puede ocurrir que lo que sea en un sentido el «mismo» mundo (las dos versiones están profundamente relacionadas) pueda ser descrito como si consistiera en «mesas y sillas» (y éstas descritas como coloreadas, poseyendo propiedades disposicionales, etc.) en una versión *y* como si consistiera en regiones espacio-temporales, partículas y campos, etc., en otras versiones. Exigir que todas ellas *deban* ser reducibles a una única versión es cometer el error de suponer que «¿cuáles son los objetos reales?» es una pregunta que tiene sentido *independientemente de nuestra elección de los conceptos*.

Lo que estoy diciendo es francamente programático. Permítanme acabar indicando brevemente a dónde conduce el programa y qué espero de él.

Muchos pensadores han argumentado que debe abandonarse la dicotomía tradicional entre el mundo «en sí mismo» y los conceptos que usamos para pensar y hablar sobre él. Por mencionar sólo los ejemplos más recientes, Davidson ha argüido que la distinción entre «esquemas» y «contenidos» no puede ser trazada; Goodman ha sostenido que la distinción entre «mundo» y «versiones» es insostenible; y Quine ha defendido una «relatividad ontológica». Como los grandes pragmatistas, estos pensadores nos han urgido a rechazar el punto de vista del espectador en metafísica y epistemología. Quine nos ha instado a aceptar la existencia de entida-

des abstractas sobre la base de que son indispensables en matemáticas,[10] y de las micropartículas y puntos espacio-temporales sobre la base de que son indispensables en física; y se pregunta ¿qué mejor justificación hay para aceptar una ontología que su indispensabilidad en nuestra práctica científica? Goodman nos ha recomendado tomar en serio las metáforas que los artistas usan a fin de reestructurar nuestros mundos, sobre la base de que son un medio indispensable para entender nuestra experiencia. Davidson, desde una posición similar, ha rechazado la idea de que el lenguaje de las actitudes proposicionales es de «segunda clase». Estos pensadores han estado, de alguna forma, dudando de si continuar extendiendo el mismo enfoque a nuestras imágenes morales de nosotros mismos y del mundo. Pero, ¿qué puede significar el abandono de la perspectiva del espectador en filosofía si no extendemos este enfoque pragmático a las «versiones» más imprescindibles que poseemos de nosotros mismos y del mundo? Como William James (y como mi maestro Morton White[11]), propongo hacer exactamente esto. En las conferencias siguientes, ilustraré el punto de vista del realismo pragmático en ética echando una mirada a algunas de nuestras imágenes morales, particularmente a las que subyacen al valor democrático central de la *igualdad*. Aunque la realidad y la verdad son viejas, y con una apariencia superficial de temas «áridos», intentaré

10. «On What There Is», recogido en *From a Logical Point of View*, Cambridge, Ma: Harvard, 1953.

11. White ha defendido que se haga esto desde hace tiempo y también recientemente (*Toward Reunion in Philosophy*, Cambridge, Ma: Harvard University Press, 1956; *What Is and What Ought to Be Done*, Oxford: Oxford University Press, 1981).

convencerles en el curso de estas conferencias de que es la persistencia de supuestos obsoletos sobre estos temas «áridos» lo que sabotea la discusión filosófica acerca de todos los temas «apasionantes», por no hablar de la posibilidad de hacer justicia a la realidad y al misterio de nuestro mundo de sentido común.

CONFERENCIA II

EL REALISMO Y LA RAZONABILIDAD[12]

Algunas cuestiones de lógica filosófica son capaces de dividir a los filósofos en campos de batalla. Desde mediados del siglo XX, esto ha ocurrido con la pregunta sobre el *status* de las afirmaciones disposicionales (y con la pregunta estrechamente relacionada del *status* de los condicionales contrafácticos). Para algunos filósofos, las disposiciones son simplemente parte del «mobiliario del universo»; para otros, el uso de una noción disposicional en el análisis filosófico es un signo de «debilidad de criterios», del deseo de «explicar lo oscuro por lo todavía más oscuro»; para otros más (quizás para la mayoría silenciosa), las nociones disposicionales son inevitables en la práctica, pero dudosas desde un punto de vista teórico. Se trata de un estado de cosas relativamente nuevo: los autores que construyeron el canon de la «filosofía moderna» (o, al menos, de la filosofía que va del siglo XVII a la mitad del siglo XIX) disponían de la noción de una potencia (es decir, una propiedad disposicional) sin ningún visible remordimiento de conciencia.

Quizás esto no sea sorprendente, ya que es sólo desde la aparición de la lógica matemática cuando nos hemos dado cuenta de lo difícil que es dar una interpretación de los condicionales contrafácticos y

12. Como se explicó en el Prefacio, esta «conferencia» no fue realmente leída en Washington.

de los predicados disposicionales en términos veritativo-funcionales.[13] Pero, en cierto sentido, debería haber sido evidente desde hace mucho tiempo que hablar de poderes en la filosofía «moderna» era problemático, ya que tal lenguaje es algo que depende de la filosofía medieval, no algo que pertenezca de pleno derecho a la nueva imagen. El corazón de la nueva imagen es la nueva concepción del mundo «externo», la concepción del mundo externo como si estuviera gobernado por leyes *estrictas* de una forma que nos es familiar desde el trabajo de Newton y sus sucesores. Es esta concepción la que motiva la división de propiedades en primarias y secundarias, o en propiedades intrínsecas de las cosas externas y potencias para afectar a la mente del observador. Un mundo gobernado por un sistema de ecuaciones diferenciales es una cosa; un mundo medieval (o aristotélico) gobernado por formas sustanciales que se manifiestan a sí mismas como «tendencias» más que como leyes sin excepciones es algo distinto. La imagen cartesiana es confusa. Exhibe ambas formas de explicación, el fisicalismo moderno y el «tendencialismo» medieval, en una coexistencia infeliz. La nueva imagen de la naturaleza –el mundo máquina– no debe tener sitio para las «tendencias» clásicas.

En la conferencia anterior, se argumentó esto

13. En lógica, se dice que una cierta forma de conectar enunciados es veritativo-funcional si el valor de verdad del enunciado resultante puede ser determinado dados sólo los valores de verdad de sus componentes. Todos los condicionales contrafácticos tienen antecedentes falsos y, normalmente, también consecuentes falsos; pero algunos de ellos son verdaderos y algunos falsos. En consecuencia, el contrafáctico no es una función-de-verdad de sus partes.

con la ayuda del ejemplo del predicado de color «rojo». Algo es rojo si muestra cierta tendencia –la tendencia a producir ciertas «impresiones sensoriales» (de acuerdo a la historia de los siglos XVII y XVIII), o cierto «estado cerebral» (una alternativa a la historia dualista que se remonta al menos hasta Diderot, si no hasta Hobbes)–, o (en una historia que es muy simple, pero, al menos, evita el problema mente-cuerpo) si tiene la tendencia a absorber selectivamente y reflejar ciertas longitudes de onda de luz. Pero, ¿qué significa «tener la tendencia»? Las tendencias, como dije en la conferencia de ayer, no ejemplifican las operaciones de las leyes estrictas (en el sentido moderno de «ley estricta»); son cosas poco sólidas que se manifiestan a sí mismas «bajo condiciones normales». Para analizar el lenguaje disposicional necesitamos un análisis de la expresión «bajo condiciones normales», o algo similar, y, de hecho, los intentos de producir una teoría hechos por autores contemporáneos[14] involucran nociones tales como la «similitud» de un mundo posible en su totalidad con otro mundo en su totalidad –nociones que intentan expresar, o al menos sustituir, la deseada noción de un estado de cosas «normal»–. Sin embargo, las nociones actualmente más de moda –la noción de «similitud» de mundos posibles– sólo ilustran la distancia del lenguaje contrafáctico (y disposicional) de la imagen del mundo de la física: lo ilustran introduciendo un primitivo metafísico que resulta como un dolor de muelas.

Otros filósofos se conforman con introducir los

14. David Lewis, *Counterfactuals*, Cambridge, Ma: Harvard University Press, 1973.

predicados disposicionales uno a uno, según se necesiten, sin ningún intento de analizar o explicar el lenguaje disposicional en general. En ocasiones, esto se puede justificar (desde un punto de vista «objetivista») mostrando que el predicado así introducido es de la misma extensión que un predicado no disposicional (quizás uno estructural). Pero muchas nociones disposicionales –es decir, «rojo», «venenoso», «inclinado a decir *da* si el lingüista dice *gavagai* y ambos están viendo un conejo»– no son, casi seguro, de la misma extensión que predicados definibles en el lenguaje de la física fundamental.

Algunos otros filósofos han sugerido que los predicados disposicionales no son, en general, el tipo de predicados para los cuales uno debería esperar que existieran condiciones necesarias y suficientes. Quizás una palabra como «venenoso» esté sólo parcialmente definida; quizás cuando encontramos una sustancia nueva que los seres humanos son capaces de ingerir, inhalar o tocar, extendemos la noción de ser venenoso justamente de la misma forma como extendemos nuestras otras nociones (incluyendo la noción de lo que es «normal») en las circunstancias dadas.[15] Otros filósofos han sugerido que las afirmaciones disposicionales como «X es venenoso» en absoluto predican una *propiedad*; son maneras en las cuales realizamos el acto de habla de *permitir una inferencia*. Como expresa J. L. Mackie en sus últimos trabajos, tales afirmaciones pueden ser defendibles bajo condiciones apropiadas

15. Esta idea estaba implícita en el tratamiento de Carnap de los predicados disposicionales vía una «reducción de sentencias» en «Testability and meaning», *Philosophy of Science*, 3: 420-468 (1936); 4: 1-40 (1937).

sin poseer ninguna propiedad que un realista reconocería como «verdad» (ellos «no son simplemente verdaderos», declaró).[16] Lo que muchas de estas teorías tienen en común es un rechazo a que la semántica de los enunciados disposicionales sea la semántica clásica bivalente de las condiciones-de-verdad. O las sentencias disposicionales no son, en absoluto, «simplemente verdaderas» ni «simplemente falsas», dicen estos autores, o, acaso, son verdaderas y falsas sólo en ciertos casos (los casos en los cuales el predicado disposicional ha sido definido), y falta que se dé un valor de verdad en todos los demás casos. (En cualquier versión de esta perspectiva, el predicado disposicional carece de una extensión bien definida.)

Como mencioné en la última conferencia, surgen alternativas similares en relación a las nociones de *causalidad* y de *explicación* (concebida como una relación entre eventos o entre «situaciones», más que como una relación entre afirmaciones). De forma parecida a las disposiciones, las relaciones causales y explicativas pueden ser estrictas (el evento o «situación» descritos como la causa pueden estar conectados por leyes estrictas al evento o situación que son considerados como el efecto) o pueden ser imprecisas (el evento o situación descritos como la causa pueden producir el efecto sólo «bajo condiciones favorables»). Y las relaciones causales imprecisas son, una vez más, embarazosas desde el punto de vista de la imagen «objetivista» –la imagen de la naturaleza como la máquina del mundo.

Si pudiéramos definir en términos fisicalistas

16. J. L. Mackie, *The Cement of the Universe*, Nueva York, Oxford University Press, 1974.

qué es para un hecho o situación ser sólo una «circunstancia acompañante», podríamos ser capaces de explicar «X produjo Y» en el sentido de que, *dadas las circunstancias concomitantes*, se seguiría de las leyes físicas que Y ocurriría si ocurriera X; pero, desafortunadamente, una distinción intrínseca entre situaciones que son capaces de ser «productivas» y situaciones que sólo son «circunstancias» acompañantes tiene más que ver con las nociones medievales (y aristotélicas) de «causalidad eficiente» que con las post-newtonianas. Y, una vez más, algunos filósofos han propuesto *rechazar* del todo las relaciones causales y explicativas imprecisas,[17] mientras que otros han sugerido que las relaciones causales y explicativas imprecisas[18] tengan sólo «condiciones de asertabilidad» y no «condiciones de verdad».

Mi propia perspectiva –la perspectiva que comencé a esbozar para ustedes en la última conferencia– se diferencia de todas éstas. Todos estos autores asumen que podemos trazar la distinción entre lo que es «simplemente verdadero» y lo que sólo tiene «condiciones de asertabilidad», o el corte entre lo que es ya verdadero o falso y lo que es una «extensión del uso previo» (aunque sea un uso que todos nosotros hacemos de la misma forma), o entre lo que es una «proyección» y lo que es una propiedad independiente y unitaria de las cosas en sí

17. Por ejemplo, Hempel propuso considerar «explicaciones» completas sólo a aquellas explicaciones que se adecuasen a su estricto modelo nomológico deductivo (véase Hempel and Oppenheim, «The Logic of Scientific Explanation», recogido en Feigl y Brodbeck (comps.), *Readings in the Philosophy of Science*, 319-352, Nueva York: Appleton-Century-Crofts, 1953).

18. Mackie se refería a la noción que usamos en ellas como nuestra «paleolítica» noción de causación. Véase n. 5.

mismas. Pienso que, al menos epistémicamente, el intento de dibujar esta distinción, de hacer este corte, ha sido un fracaso total. Ha llegado el momento de ensayar la hipótesis metodológica de que no puede hacerse ningún corte como ése.

Recuerdo una conversación con Noam Chomsky, hace muchos años, en la cual él insinuaba que los filósofos a menudo toman continuos perfectamente sensatos y tienen problemas al intentar convertirlos en dicotomías. Considérese, por ejemplo, el continuo entre lo relativamente «subjetivo» (o, al menos, relativo a una cultura y a unos intereses) y lo relativamente «objetivo» (o, al menos, independiente de una cultura y de unos intereses). Prefilosóficamente, muchos de nosotros probablemente estaríamos de acuerdo en el orden que las siguientes propiedades tendrían a través de este continuo:

1. Ser muy divertido (como en «la conducta de los niños pequeños es a menudo muy divertida»).
2. Ser una región del espacio que contiene al menos un átomo de hidrógeno (adóptese para esto la física clásica –¡por favor!, no, la mecánica cuántica o la relatividad).
3. Ser soluble.
4. Un caso aislado de condicional contrafáctico –p. ej., la propiedad que predicamos de una cerilla particular en un momento particular cuando decimos de ella que se *habría* encendido si *hubiera* sido frotada en ese momento.
5. Significar «¿habla francés?» (predicado de una expresión concreta).

Supongo que una persona normal podría ordenar estas predicaciones como sigue (tomando el fi-

nal izquierdo de la línea para representar lo «subjetivo» y el final derecho para representar lo «objetivo»):

Ser divertido	Contra-fáctico	Significar «...»	Ser soluble	Contiene hidrógeno

(Una ordenación subjetivo-objetivo admisible)

Ahora, tan pronto como se nos pide hacer un «corte tipo Dedekind» –convertir esta ordenación en una dicotomía–, encontramos que no hay para nada un acuerdo en nuestras intuiciones filosóficas. Quine, por ejemplo, pondría el corte entre 5 y 3, considerando «objetivos» tanto a los predicados disposicionales (tales como «soluble») como a los predicados no disposicionales de la física fundamental y, más o menos, subjetivos (o de «segunda clase», en su terminología) a todos los demás. Algunos filósofos pueden discrepar conmigo en la posición de 5 acerca de la asignación de significado –algunos la considerarán más «objetiva» que la asignación de *solubilidad* a una sustancia– y dibujarán una línea después de 1, 4 y 3. Los filósofos que se sientan «cómodos» con los contrafácticos aún harán otra elección respecto a la situación del «corte», colocándolo inmediatamente después de 1 –es decir, contando «divertido» como subjetivo y todo lo demás como «objetivo»–. Pero según mi propia perspectiva, como dije (y quizás también la de Chomsky, si le entendí bien), la empresa no vale la pena. El juego se ha acabado. Podemos hacer una ordenación a grandes rasgos (aunque incluso aquí hay desacuerdos), pero la idea de «un punto en el cual» la subjetividad cesa y la Objetividad-con-una-

O mayúscula comienza se ha mostrado quimérica.

Si esto es correcto, entonces deben ser abandonadas otras famosas dicotomías. Dos de ellas ya han sido mencionadas, son:

Proyección/Propiedad de la cosa en sí misma
y
«Potencia»/Propiedad de la cosa en sí misma.

El rechazo de estas tres dicotomías es la esencia del «realismo interno» que defendí ante esta misma concurrencia hace nueve años.

Mi rechazo de estas dicotomías perturbará a muchos, como ha de ser. Sin la constricción de intentar «salvar las apariencias», la filosofía se convierte en un juego en el cual cualquiera puede –como en general se hace– opinar justamente sobre *cualquier cosa*. A menos que tomemos seriamente nuestras intuiciones, no podremos hacer en absoluto filosofía *seria*. En consecuencia, respeto a los filósofos que insisten en que las dicotomías tradicionales son profundamente intuitivas, y que «necesitan mucho convencimiento» antes de abandonarlas.

Pero si la filosofía que simplemente desprecia nuestras intuiciones no vale la pena, la filosofía que intenta preservarlas *todas* se convierte en un vano intento de volver al pasado. Hay fenómenos que realmente desafían nuestras intuiciones: el fenómeno que Husserl describió en *Crisis of the European Sciences*, el fracaso del gran proyecto del siglo XVII de intentar convertir la física en metafísica «objetivismo» –fracaso que describí en la anterior conferencia–, es uno de ellos. Por un lado, la ciencia del siglo XVII tuvo éxito en hacer añicos los fundamentos medievales del conocimiento –y no sólo del co-

nocimiento, sino también de la religión, la política y la moralidad–. Por otro lado, la corriente de pensamiento que decía «bien, si la ciencia hace añicos todo esto, bienvenida sea. La ciencia nos dará en su lugar algo mejor» parece ahora agotada. (Ya le parecía agotada a Kant, y no porque Kant fuera un enemigo de la ciencia de la Ilustración; por el contrario, fue un gran científico y un gran hombre de la Ilustración.) La ciencia es maravillosa a la hora de destruir respuestas metafísicas, pero es incapaz de proporcionar sustitutos. La ciencia se lleva los fundamentos sin proporcionar un reemplazo. Queramos estar allí o no, la ciencia nos ha colocado en la situación de tener que vivir sin fundamentos. Fue sorprendente cuando Nietzsche lo dijo, pero hoy día es un lugar común; *nuestra* posición histórica –y no hay ningún final a la vista– es la de tener que filosofar sin «fundamentos».

La imposibilidad de imaginar a qué se parecerían unos «fundamentos» creíbles es un fenómeno, pero no el único, que desafían nuestras «intuiciones». Desde finales del siglo XIX, la ciencia misma ha comenzado a tomar una apariencia «no clásica» –esto es, una apariencia que no es la del siglo XVII–. En la última conferencia describí el fenómeno de la relatividad conceptual, el cual tiene ilustraciones simples, como las que usé, pero que ha resultado omnipresente en la ciencia contemporánea. Que hay formas de describir lo que son (en algún sentido) los «mismos hechos» que son (en algún sentido) «equivalentes» pero, también, (en algún sentido) «incompatibles» es un notable fenómeno no clásico. Todavía los lógicos contemporáneos y los teóricos del significado generalmente teorizan como si no existiera. Si pretender abandonar *todas* nuestras

«intuiciones» es un mero espectáculo, mantenerlas todas ellas requeriría de nosotros filosofar como si el fenómeno que justamente acabo de recordarles no existiera. La tarea del filósofo, tal como la veo, consiste en discernir *cuáles* de nuestras intuiciones podemos mantener responsablemente y cuáles debemos abandonar en un período de cambio conceptual, así como material, enorme y sin precedentes.

Así pues, si rechazo las dicotomías que acabo de mostrar, no es porque no reconozca su atractivo intuitivo, o porque ese atractivo intuitivo no me parezca gran cosa. Es, más bien, porque estas dicotomías se han convertido en lentes distorsionadoras que nos impiden ver los fenómenos reales –los fenómenos que he estado describiendo– en su completa extensión y significado.

Aún me llamo a mí mismo un «realista» –aunque lo deletree en minúsculas–, pero ¿se *puede* ser un realista sin las dicotomías? En particular, ¿no es la dicotomía entre lo que es una «proyección humana» –lo que no es «simplemente verdadero», lo que tiene «condiciones de asertabilidad» más que «condiciones de verdad realista»– y lo que está en las cosas «mismas» *constitutiva* del realismo?

Parte de mi respuesta a esta pregunta fue dada en la primera conferencia. Lejos de ser constitutiva del realismo de *sentido común*, esta dicotomía tiende a socavarlo, como intenté mostrar. Pero otra parte de la respuesta debe consistir en mostrar que el rechazo de esta dicotomía no es una simple capitulación a cierta variedad de relativismo cultural, o a la idea de que todo esquema conceptual es tan bueno como cualquier otro.

Lo que es extraño, en relación al temor de que sólo el realista metafísico pueda salvar al sentido

FILOSOFIA
Y LETRAS

común del demonio del relativismo, es que incluso los realistas metafísicos reconocen que los mandatos de la racionalidad van más allá que lo que ellos quieren llamar «verdad realista». Mackie no pensó que las afirmaciones causales del lenguaje ordinario, por ejemplo, «el fallo de la válvula de seguridad causó que la caldera explotara», fueran «simplemente verdaderas», aunque, ciertamente, él habría distinguido entre las «razonables» y las «no razonables». Quizás tales afirmaciones tienen sólo condiciones de «asertabilidad» más que condiciones de «verdad», quizás se utilizan para permitir «licencias inferenciales» más que para «describir», pero esto no las hace arbitrarias. Si nos permitimos esperar que X se disuelva al ser puesto en agua, cuando X es un terrón de azúcar, esto es parte de una práctica cuyo éxito podemos explicar; y si nos permitimos la misma licencia cuando X es un trozo de acero, la naturaleza nos mostrará nuestro error. De la misma forma, Quine niega que «X significa *¿hablas francés?*» formule un «hecho», aunque X sea la conocida frase francesa *parlez-vous français?*; pero seguramente él respondería a la pregunta «¿qué quiere decir *parlez-vous français*?» con «significa *¿hablas francés?*» y no con «significa *cochero, pare, la carretera está muy mal: ¡mire fuera!, perderá el pavo*». Que una respuesta a este tipo de pregunta tenga un valor «heurístico», y la otra no, es algo que él mismo señala. (No estoy pretendiendo que Quine sea un «realista metafísico» según mi punto de vista, ya que él no acepta la teoría de la verdad como correspondencia; aunque su «robusto realismo» tiene una importante característica en común con el realismo metafísico –a saber, la existencia de una barrera tajante entre aquello sobre lo que hay «hechos decisi-

vos» y aquello que tiene sólo un valor «heurístico», o un valor cuando nuestros intereses no son sólo «teóricos».)

En resumen, mi propia posición aún involucra el rechazo de otra dicotomía:

(Tipo de afirmación)

Poseyendo sólo condiciones de asertabilidad	VS.	Poseyendo condiciones de verdad

Podemos saber que es «verdad» simplemente, que el agua habría hervido si hubiese encendido el fuego, sin tener la más ligera idea sobre si esta «verdad» es una «verdad realista» (algo «simplemente verdadero» en el sentido de Mackie) o sólo una idealización de «asertabilidad garantizada». Ni necesitamos suponer que la pregunta tiene sentido. Rechazar la dicotomía *entre* clases de «verdad» –clases de verdad en el mundo del sentido común– no es lo mismo que decir que «todo vale».

Realidad sin dicotomías

¿Cómo puede uno estar seguro de que esto no sea puro idealismo lingüístico? Quizás el mejor sitio para empezar sea la explicación que di del realismo interno en la primera conferencia. Esa explicación ciertamente parece un «idealismo lingüístico»; según dije, cuántos objetos hay en el mundo (y aunque ciertos objetos –puntos espacio-temporales individuales, en el segundo ejemplo que utilicé– existan completamente como «particulares» individua-

les) es algo relativo a la elección de un esquema conceptual. ¿Cómo puede uno proponer esta clase de doctrina relativista y todavía pretender creer que hay algo en la idea de una «externalidad», algo en la idea de que «allí fuera» hay algo independiente del lenguaje y de la mente?

Bien, realmente no es tan difícil. Vean de nuevo el cuadro que les mostré:

Mundo 1	Mundo 2
x1, x2, x3	x1, x2, x3, x1 + x2, x1 + x3, x2 + x3, x1 + x2 + x3
(Un mundo al estilo de Carnap)	(El «mismo» mundo al estilo de los lógicos polacos)

Cómo vamos a responder a la pregunta «¿cuántos objetos hay?» –el método de «contar» o la noción de lo que constituye un «objeto»– depende de nuestra elección (llamemos a esto una «convención»); pero la *respuesta* no por eso se convierte en un asunto de convención. Si escojo el lenguaje de Carnap, debo decir que hay tres objetos porque *éstos son cuantos objetos hay*. Si escojo el lenguaje del lógico polaco (esto es, el lenguaje de un lógico polaco que todavía no ha inventado, recuerden, el «objeto nulo» O), debo decir que hay siete objetos, porque *éstos son cuantos objetos hay* (en el sentido de «objeto» del lógico polaco). Hay «hechos externos», y podemos *decir cuáles son*. Lo que no podemos decir –porque no tiene sentido– es que los hechos sean *independientes de todas las elecciones conceptuales*.

Una metáfora que se emplea a menudo para expresar esto es la metáfora del «molde de pastas».

Las cosas independientes de todas las elecciones conceptuales son la masa; nuestra contribución conceptual es la forma del molde. Desafortunadamente, esta metáfora no ayuda realmente a entender el fenómeno de la relatividad conceptual. Tómenla seriamente y se verán forzados de nuevo a responder a la cuestión «¿cuáles son las distintas partes de la masa?». Si responden que (en el presente caso) los «átomos» de la masa son x1, x2, x3 y que las otras partes son las sumas mereológicas que contienen más de un «átomo», entonces habrán *adoptado* simplemente la versión del lógico polaco. Insistir en que ésta es la perspectiva correcta de la situación metafísica es sólo otra forma de insistir en que las sumas mereológicas *realmente* existen. Pero el realismo interno niega que ésta sea una manera *más* «correcta» de ver la situación que la de insistir en que sólo existen realmente los «individuos» de Carnap. El metafísico que adopta la última perspectiva puede explicar también el éxito de la versión del lógico polaco; después de todo, puede decir que cuando el lógico polaco afirma, como podría hacerlo, que

(I) Hay al menos un objeto que es parcialmente rojo y parcialmente negro

esto tiene que ser entendido como una útil *façon de parler* más que como algo «literalmente verdadero». Bajo un adecuado esquema de traducción (y tal esquema puede ser dado fácilmente de manera recursiva, en el caso del tipo de lenguajes de primer orden que Carnap tenía en mente en relación con estos ejemplos sencillos), I pasaría a expresar nada más lo que

(II) Hay al menos un objeto rojo y hay al menos un objeto negro

expresa cuando es escrito en el lenguaje carnapiano. (Para verificar esto, asumiendo que «rojo» y «negro» sean predicados del lenguaje de Carnap, observen que la única forma en que un objeto del lógico polaco –una suma mereológica– puede ser parcialmente rojo es conteniendo un átomo rojo, y que la única forma en que puede ser parcialmente negro es conteniendo un átomo negro. En consecuencia, si I es verdad en el lenguaje del lógico polaco, entonces hay al menos un átomo rojo y al menos un átomo negro –lo cual es lo que II expresa en el lenguaje de Carnap–. A la inversa, si hay al menos un átomo negro y al menos un átomo rojo, entonces su suma mereológica es un «objeto» –en el sentido del lógico polaco– que es parcialmente rojo y parcialmente negro.) Pretender que tal esquema de traducción muestra lo que «realmente está pasando» sólo es una forma de insistir en que las sumas mereológicas *no* «existen realmente».

La metáfora del molde de pastas *niega* (más que explica) el fenómeno de la relatividad conceptual. La otra manera de encarar nuestro pequeño ejemplo –produciendo un esquema de traducción que *reinterpreta las conectivas lógicas (en este caso, la existencia)* de tal manera que cada afirmación hecha en el lenguaje «más rico» pueda ser «traducido» al lenguaje más «parsimonioso»– puede, también, ser empleada para negar el fenómeno de la relatividad conceptual; pero, en cualquier caso, es más sofisticada que la metáfora del molde de pastas. La metáfora del molde de pastas da por supuesto que todas las afirmaciones que consideramos verdade-

ras en nuestras múltiples versiones realmente son verdaderas; sólo que las variables de la cuantificación recogen diferentes sumas mereológicas como argumentos en el caso de los distintos lenguajes. La estratagema de la *reinterpretación* va más allá de esto reconociendo que lo que para una persona es una afirmación de «existencia» para otra persona puede ser alguna otra cosa.

A veces se sugiere que en tales casos *no* deberíamos ser neutrales, que deberíamos siempre adoptar la versión más parsimoniosa. «Si no tenemos que postular objetos discontinuos tan extraños como las sumas mereológicas, entonces ¿no tendríamos que tomar esto como una razón para concluir que no existen realmente, que son sólo (como mucho) una *façon de parler*?»

Para este movimiento metafísico hay, inevitablemente, una contestación igualmente metafísica: «¿No son casi todos los "objetos" de los que hablamos –sillas y mesas, nuestros propios cuerpos, países, por no mencionar objetos científicos tales como sistemas solares y galaxias– "objetos discontinuos extraños"? Difícilmente se sigue que no existan realmente. Pero, si mi cuerpo existe, si esta silla existe, si existe el sistema solar, entonces, ¿por qué no deberíamos decir que el objeto discontinuo formado por *mi nariz y la Torre Eiffel* también existe? Se trata de un objeto poco natural, seguro, pero ¿qué tiene que ver la "naturalidad" de un objeto con su *existencia*?».

Lo que es correcto en el segundo de los caminos que hemos considerado para reconciliar las dos versiones o «mundos» –reinterpretar el cuantificador existencial– es que las nociones de «objeto» y «existencia» no se tratan como sacrosantos, como si tu-

vieran sólo un uso posible. Es muy importante reconocer que el cuantificador existencial mismo puede utilizarse de diferentes formas –formas en consonancia con las reglas de la lógica formal–. Lo que estaría mal, si lo hiciéramos, sería aceptar esta idea, y después acabar eligiendo *un* uso del cuantificador existencial –el uso en la versión de Carnap– como el único metafísicamente *serio*. Pero demos un paso más. Considérese la posición de que uno puede *o* tratar la versión de Carnap como «correcta» e interpretar la versión del lógico polaco como una *façon de parler* a la manera ilustrada por la reinterpretación de I como II, *o* tratar la versión del lógico polaco como «correcta» e interpretar la versión de Carnap como un lenguaje en el que el dominio de las variables individuales se restringe a átomos (como se sugiere en la metáfora del molde de pastas). Esto es, tómese la posición en la que uno estaría «en lo correcto» en cualquiera de los casos por igual. Entonces, ¡habrán llegado a la posición que he llamado «realismo interno»!

Lo que es erróneo con la noción de objetos que existen «independientemente» de los esquemas conceptuales es el hecho de no haber estándares ni siquiera para el uso de las nociones lógicas al margen de las elecciones conceptuales. Lo que la metáfora del molde de pastas intenta conservar es la idea ingenua de que al menos una categoría –la antigua categoría de objeto o sustancia– tiene una interpretación absoluta. La alternativa a esta idea no es la concepción de que, en algún sentido inconcebible, todo no es *más* que lenguaje. Podemos y debemos insistir en que algunos hechos están ahí para ser descubiertos y no para ser legislados por nosotros, aunque eso sea algo que hay que decir una vez que

se ha adoptado una manera de hablar, un lenguaje, un «esquema conceptual». Hablar de los «hechos» sin especificar el lenguaje que va a utilizarse no es hablar de nada; la palabra «hecho» no tiene fijado su uso por la realidad misma más de lo que lo tiene la palabra «existe» o la palabra «objeto».

Por supuesto, la adopción del realismo interno es la renuncia a la noción de «cosa en sí». Y aquí está la conexión entre el ejemplo casi trivial que hemos discutido y las dicotomías metafísicas profundas que discutimos con anterioridad. El realismo interno dice que la noción de una «cosa en sí» no tiene sentido; y *no* porque «no podamos conocer las cosas en sí». Ésta fue la razón de Kant, pero éste, aunque admitía que la noción de una cosa en sí *puede ser* «vacía», le permitió a ésta tener todavía un sentido de tipo formal. El realismo interno dice que no sabemos de qué estamos hablando cuando hablamos sobre las «cosas en sí». Y esto significa que la dicotomía entre propiedades «intrínsecas» y propiedades que no son intrínsecas también colapsa –se colapsa porque se supone que las propiedades «intrínsecas» son sólo las propiedades que las cosas tienen «en sí mismas»–. La cosa en sí y la propiedad que la cosa tiene «en sí misma» pertenecen al mismo círculo de ideas, y ya es el momento de admitir que lo que el círculo abarca es un territorio sin valor.

Una dicotomía cuya relación con estas nociones puede no ser tan evidente es la dicotomía entre una «semántica de condiciones de verdad» y una «semántica de condiciones de asertabilidad». Con todo, ¿qué podría fundamentar la pretensión de que ciertos tipos de afirmaciones, por ejemplo, «si *hubiera* puesto un cazo de agua en la cocina, y encendido el fuego, el agua *habría* hervido», sólo tienen «condicio-

nes de asertabilidad» y no «condiciones de verdad»? ¿Qué, aparte de una idea preconcebida de lo que es y no es «ontológicamente extraño», es decir, de lo que es y no es capaz de ser parte del mundo como el mundo es «en sí mismo»? Como argumenté en la conferencia de ayer, el problema con esta idea preconcebida, en su versión tanto humeana como cartesiana, era su incapacidad para contarnos nada sobre la mente (o, si lo prefieren, sobre la «intencionalidad») que no estuviera plagado de contradicciones y cargado de supuestos arbitrarios y poco convincentes; y yo argüí que éste sigue siendo hoy su problema.

¿Cómo es el mundo sin dicotomías? Parece, a la vez, familiar y diferente. Parece familiar en tanto no intentamos por más tiempo dividir la realidad mundana en una «imagen científica» y una «imagen manifiesta» (o nuestra doctrina en fase de desarrollo en un sistema conceptual de «primera clase» y otro de «segunda clase»). Las mesas y las sillas (sí, y los cubitos de hielo rosa) existen tanto como los quarks y los campos gravitatorios, y el hecho de que este cazo de agua habría hervido si lo hubiera puesto en la cocina, y hubiera encendido el fuego, es tanto un «hecho» como lo es la circunstancia de que el agua pese más de 200 gr. La idea de que la mayor parte de nuestra realidad mundana es una ilusión (una idea que ha obsesionado a la filosofía occidental desde Platón, a pesar del valiente contraataque de Aristóteles) se abandona de una vez por todas. Pero la realidad mundana parecerá diferente porque estamos forzados a reconocer que muchas de nuestras descripciones corrientes reflejan nuestros intereses y elecciones.

Imaginen que la válvula de escape de una olla a

presión se atasca y que la olla a presión explota. Decimos –y el relativista conceptual lo verá como una afirmación perfectamente «verdadera», sin hacer ninguna protesta sobre si es «simplemente verdadera» o sólo una «buena licencia inferencial»– «la válvula atascada causó que la olla a presión explotara». No decimos «la presencia de ▲ causó que la olla a presión explotara», donde ▲ es, digamos, una parte arbitraria de la superficie de la olla con una forma irregular, de 0,1 cm de área. Con todo, en la física de la explosión, el papel que desempeña la válvula atascada es exactamente el mismo que el papel de ▲: la ausencia de cualquiera de ellos habría permitido al vapor escapar, bajando la presión y evitando la explosión.

¿Por qué, entonces, hablamos de una de estas cosas y no de la otra como «causante» de la explosión? Bien, sabemos que la válvula «debería haber» permitido salir al vapor –ésta es su función, aquello para lo que fue diseñada–. Por otro lado, el elemento en la superficie ▲ no estaba haciendo nada «incorrecto» al impedir al vapor escapar: contener el vapor es la «función» de la superficie de la cual ▲ es parte. En consecuencia, cuando preguntamos «¿por qué tuvo lugar la explosión?», sabiendo lo que sabemos y teniendo los intereses que tenemos, nuestro «espacio de explicación» consta de las alternativas:

1. La explosión tiene lugar.
2. Todo funciona como es debido.

Lo que queremos saber, en otras palabras, es por qué 1 es lo que ocurrió, *en oposición a* 2. Simplemente no estamos interesados en por qué 1 es lo que ocurrió *en oposición a* alternativas tales como:

3. El elemento en la superficie ▲ falta, y no tiene lugar ninguna explosión.

Esta «relatividad explicativa» es paralela a la relatividad en nuestro uso de locuciones tales como «causó» y «la causa». Ya que la pregunta «¿por qué explotó la olla a presión?» presupone un espacio de explicación que no incluye la alternativa 3, o alternativas similares, y entendemos tales factores como la presencia de ▲ como «condiciones del entorno» y no como «causas».

Esta relatividad de las causas a los intereses y a las condiciones del entorno, no mencionada en la explicación de «ciencia dura» del suceso en cuestión, no convierte a la causalidad en algo que nosotros simplemente legislemos. Dados nuestros intereses y lo que vemos como las condiciones conexas del entorno, sería simplemente falso decir que fue la pared de la olla a presión lo que causó la explosión (a menos que aquélla fuese defectuosa, y que fuera ese defecto, y no el estado de la válvula, lo que «explicara» la explosión). Nuestro esquema conceptual restringe el «espacio» de descripciones que tenemos a nuestra disposición; pero no predetermina las respuestas a nuestras preguntas.

Es comprensible, sin embargo, que muchos filósofos sacaran una moraleja diferente de nuestra historia. ¿No conduce la situación misma de modo natural a una dicotomía? ¿No deberíamos ver la descripción que la «ciencia dura» hace de la situación («El incremento de presión en el recipiente cerrado hasta que se sobrepasa cierto coeficiente. El material roto a continuación...»), con sus leyes exactas y coeficientes numéricos, como la descripción de los «hechos objetivos», y ver en la selección del trozo de material –o lo que sea– que impidió a la válvula fun-

cionar como «la causa» un pensamiento semimágico de la Edad de Piedra? Si queremos ser generosos y dejar un sitio a esta útil manera de hablar, a la vez que negar que exista una distinción entre «causas» y «condiciones del entorno» en la naturaleza misma, sólo podemos decir que las afirmaciones causales tienen condiciones de «asertabilidad» en el lenguaje ordinario pero no, estrictamente hablando, «condiciones de verdad».

El problema con todo esto –el problema que discutí en la primera conferencia– es que si la distinción entre causas/condiciones del entorno es fundamentalmente subjetiva, no descriptiva del mundo en sí mismo, entonces las explicaciones filosóficas habituales de la naturaleza metafísica de la *referencia* están en bancarrota. Barwise y Perry, por ejemplo, nos dicen que lo que vincula ciertos estados de hechos con ciertos estados mentales es que los estados de hechos *causan* aquellos estados mentales; en esto consiste el vínculo intencional, al menos en ciertos casos básicos metafísicos. Glymour y Devitt (de forma independiente) nos dicen que las palabras están conectadas con sus referentes a través de «conexiones causales». Richard Boyd nos dice que la «teoría causal de la referencia es correcta porque la teoría causal del conocimiento es correcta». Pero las nociones de las que dependen las teorías causales del conocimiento y la referencia –la diferencia entre una causa y una mera condición del entorno, la legitimidad de los contrafácticos– son precisamente las que la interpretación de las afirmaciones causales y de los contrafácticos como «licencias inferenciales» ponen en entredicho. Si se «salvan» estas nociones sólo en la medida en que se traten como heurísticas (como «proyecciones» en la

terminología de la primera conferencia), entonces no puede mantenerse también que ellas explican cómo llega a existir la referencia en el mundo tal como el mundo es «en sí mismo».

Tampoco ayudaría el dualismo, si quisiéramos adoptarlo. Porque, ¿qué descripción tenemos de la mente «en sí misma»? La exposición de Kant de la bancarrota de la «psicología racional» todavía se mantiene.

Más que sucumbir a la tentación de repetir palabra por palabra todas las propuestas de los siglos XVII y XVIII, tenemos que reconocer que tales afirmaciones familiares, como la afirmación de que la válvula atascada causó la explosión de la olla a presión, reflejan tanto cómo son las cosas así cuanto nuestros intereses y suposiciones sobre cómo las cosas son, y ello *sin* caer en la tentación de suponer que la descripción filosóficamente adecuada de «la manera como son las cosas» no es *otra* que «la válvula se atascó y causó que la olla a presión explotara» (o cualquier otro ejemplo). *Dado* un lenguaje, podemos describir, en un sentido «trivial», los «hechos» que hacen verdaderas y falsas las sentencias de ese lenguaje –usando las sentencias de ese mismo lenguaje–; pero el sueño de encontrar una relación universal bien definida entre una (supuesta) totalidad de *todos* los hechos y una sentencia verdadera arbitraria en un lenguaje arbitrario es sólo el sueño de una noción absoluta de un hecho (o de un «objeto») y de una relación absoluta entre sentencias y los hechos (o los objetos) «en sí»; el mismo sueño cuya imposibilidad esperé mostrar con la ayuda de mi pequeño ejemplo que involucraba a tres individuos carnapianos y siete sumas mereológicas no vacías.

CONFERENCIA III

LA IGUALDAD Y NUESTRA IMAGEN MORAL DEL MUNDO

En la conferencia precedente, intenté explicar y defender una posición que combina elementos del «realismo» con elementos del «antirrealismo». Muchos han percibido que mi posición pertenece a la tradición kantiana (entendida en sentido amplio). Y puede estar bien decir algo sobre las relaciones entre mis opiniones y el trabajo de Kant.

Hay interpretaciones de Kant bajo las que tal relación no es nada evidente, al igual que hay algunas interpretaciones de Kant bajo las que la relación de lo que voy a decir con la filosofía moral de Kant no será en absoluto evidente. Y el fallo no está enteramente en los lectores de Kant. Éste tiene, en cierto modo, dos filosofías. A veces Kant dice que la noción de un *Ding an sich*, de una «cosa en sí misma», puede ser vacía: una interpretación de esto en lenguaje contemporáneo (una interpretación seguramente controvertida) podría consistir en decir que mientras que los pensamientos sobre cómo son las cosas «en sí» pueden estar sintácticamente bien formados, y mientras que quizás tengamos una propensión natural a comprometernos con tales pensamientos, les falta cualquier inteligibilidad real. Pienso que casi toda la *Crítica de la razón pura* es compatible con una lectura en la cual uno no se comprometa con un mundo noumenal, ni, incluso, como ya dije, con la inteligibilidad de los pensamientos acerca de los noumena.

Kant dio argumentos muy fuertes en favor de la concepción que acabamos de describir, la concepción de que realmente no podemos formarnos ninguna noción inteligible de una cosa noumenal. Con todo, cuando Kant pasó a escribir su filosofía moral, postuló una «necesidad de la razón práctica pura» que requiere que creamos que la idea de los «noumena» tiene *algún* tipo de sentido, aunque seamos incapaces de decir de qué clase de sentido se trata, y a cuyo servicio debemos colocar ciertas cosas particulares afirmadas sobre el mundo noumenal: que somos noumenalmente libres; que después de la muerte nos acercaremos, de alguna manera, al *Summum Bonum*; y que hay alguna realidad noumenal que merece el nombre «Dios» (aunque Kant dice con cautela que Dios puede no ser una persona, que ésta puede ser sólo la manera que tenemos que pensar en Él en esta vida). En esta imagen –la imagen que debemos aceptar para satisfacer las «necesidades de la razón práctica»– hay dos mundos. En un mundo, el «mundo de la experiencia», todos los puntos señalados por Kant cuando él adopta la posición que describí en primer lugar son válidos, y el conocimiento científico –que es el único *conocimiento* que, de acuerdo con Kant, se llama así con propiedad– no puede ir más allá de este mundo. En el otro mundo, el mundo detrás del velo, hay un Dios, libertad e inmortalidad. Pero, obviamente, no puede haber allí *sólo* Dios, libertad e inmortalidad.

Rechazo, como muchos filósofos han rechazado, este dualismo de un mundo noumenal y uno fenoménico.[19] Pero ésta no es –repitámoslo, *no*– la úni-

19. La imagen presentada en la segunda Crítica (y a veces, desafortunadamente, en la primera) es, en muchos sentidos, un

ca filosofía de Kant. Me parece que el empuje real de la *Crítica de la razón pura* es bastante diferente. Lo que condujo a Kant a mantener un conjunto doble de libros no fue, pienso, su concepción de la verdad, sino sus ideas erróneas sobre la filosofía moral, específicamente la idea equivocada de que la filosofía moral es imposible sin garantías trascendentales que pueden ser dadas sólo si postulamos un reino noumenal.

Si la *Crítica de la razón pura* de Kant fracasa, en este sentido, como moralidad, me parece que también fracasa como religión. Para la persona religiosa, el lugar de Dios en la moralidad no debería ser primariamente repartir recompensas y castigos, como ocurre (algunas veces) en el esquema de Kant. Ni hace ningún servicio a la religión pensar en Dios como en un objeto detrás del velo –un velo tan espeso que el mismo hecho de concebir algo detrás de él debe ponerse en duda–. Pero nuestra relación con lo divino no es el tema de estas conferencias. Mi tema es más bien la pista proporcionada por la interpretación de la *primera* Crítica, para la cual la filosofía de Kant no se compromete con la distinción entre cosas en sí y proyecciones, sino que más bien sugiere un rechazo de la misma. Bajo tal interpretación, las «cosas para nosotros» de la primera Crítica son simplemente *cosas*, no «proyec-

---

salto hacia atrás a la imagen de los *Prolegomena*, en la cual el mundo noumenal simplemente *es* el mundo, el mundo tal como es «en sí mismo», y el mundo de la experiencia es el mundo de las «cualidades secundarias» de Locke –o, mejor, en lo que se convertiría ese mundo si dijéramos que la extensión y, realmente, todas las cualidades que conocemos por medios científicos o perceptuales ordinarios son «secundarias» en el sentido de Locke.

ciones», y el punto central de la primera Crítica *no es* que todo lo que nosotros conozcamos sean «experiencias» en el sentido empirista, esto es, que todo lo que nosotros conozcamos sean datos de los sentidos. La primera Crítica, tal como yo la leería, *rechaza* la noción de un dato de los sentidos, es decir, de un objeto cuya «esencia sea ser percibido» y en cuya constitución el sistema conceptual no juegue ningún papel. Para Kant, las sensaciones –los «objetos del sentido interno»– están al mismo nivel que los llamados «objetos externos». Están tan cogidas en la red de la creencia y la conceptualización como lo están los objetos externos. No representan un «dato» incorrupto, que de alguna forma *asegura* nuestro conocimiento. No entraré en más detalles porque ésta no es una conferencia histórica. Sólo les recordaré que en *Razón, verdad e historia* presenté a Kant como el primer filósofo que rechazó la idea de la verdad como correspondencia con una realidad pre-estructurada.

Si Kant *estaba* diciendo que la verdad no debe ser pensada como correspondencia con una realidad pre-estructurada o auto-estructurada, si estaba diciendo que nuestra contribución conceptual no puede ser aislada y que lo «que-hace-la-verdad» y lo «que-hace-la-verificación» de nuestras creencias está *dentro* y no *fuera* de nuestro sistema conceptual, entonces Kant puede ser llamado acertadamente el primer «realista interno». Es verdad que no acepta la clase de relatividad conceptual que yo defendí en mi primera conferencia. Por el contrario, piensa que tenemos exactamente una versión *científica* del mundo. Con todo, me pregunto si no hay una *insinuación* de la relatividad conceptual en

el hecho de que en cada Crítica –no sólo en las dos primeras– se nos presenta un *tipo* diferente de razón, y la que podría ser llamada una imagen diferente del mundo junto con cada tipo de razón: razón científica, razón ética, razón estética, razón jurídica. Aunque, de nuevo, ésta no es una conferencia histórica. Estoy proponiendo a Kant como antecesor, no tratándolo como una escritura.

Acabo de criticar la filosofía moral de Kant, porque, si es en algún sitio, es allí donde encontramos el conjunto doble de libros, la idea de un mundo detrás de un velo y un mundo para nosotros, como en la imagen corriente del «kantismo». Con todo, quiero decir que incluso en la segunda Crítica hay una profunda tensión. La segunda Crítica, más que ninguna otra, defiende la imagen que yo objeto. Sin embargo, sostengo que presenta ideas que pueden separarse de esa imagen, ideas que pueden ser el comienzo de una clase de «realismo interno» en filosofía moral. Son esas ideas, y, más ampliamente, la idea de un «realismo interno» en filosofía moral, las que quiero esbozar en las conferencias que quedan de esta serie.

Quiero indicar que la filosofía moral de Kant representa, ante todo, un replanteamiento de los valores que Kant tomó de Rousseau (y de los ideales fundamentales de la Revolución francesa), en particular del valor de la igualdad. El valor de la igualdad es, quizás, una contribución única de la religión judía a la cultura occidental. La ética griega, como la conocemos en Platón y en Aristóteles e incluso en el período helenístico, no tiene ninguna noción de igualdad humana universal. La idea de la igualdad aparece en la Biblia judía como la idea de que todos los seres humanos son creados a imagen

de Dios. Esto se conecta con algunas características del código legal judío, por ejemplo, el hecho de que la vida de un israelita es tan valiosa como la vida de cualquier otro israelita, que el ojo de un israelita es tan valioso como el ojo de cualquier otro israelita, etc. Lo que convierte a este punto de vista en una innovación radical cuando se compara con el Código de Hammurabi y con otros códigos antiguos es que estos últimos aplicaron la idea de «igualdad en las penas» sólo a los iguales sociales: un noble que mataba a un esclavo sólo tenía que pagar una multa. La idea de un conjunto de penas que debería ignorar la clase social, la idea de una justicia que, en el lenguaje bíblico, no «respeta a las personas», es decir, que se muestra parcial, fue una concreción de la igualdad.

Más tarde, por supuesto, la idea de la igualdad fue separada de sus raíces específicamente religiosas; un efecto de esta separación, un efecto que vemos hasta el día presente, es que la idea de la igualdad se convierte en algo *misterioso* y, por esta razón, expuesta a burlas. Me pregunto a veces cuánta gente sigue creyendo realmente en la igualdad humana, o sólo dicen lo que dicen desde un cierto sentimentalismo. Quizás los siguientes dos principios, vagos como son, puedan capturar el contenido mínimo que la idea de igualdad, que la cultura occidental toma de la Biblia, parece tener:

I. Hay algo en los seres humanos, algún aspecto que tiene un significado moral incomparable, con respecto a lo cual todos los seres humanos son iguales, no importa lo desiguales que puedan ser en talento, logros, contribución social, etc.

II. Incluso aquellos con menos talento, o cuyos logros sean menores, o cuya contribución a la sociedad sea más pequeña, son merecedores de respeto.[20]

Estos principios dicen que deberíamos mostrarnos unos a otros un *respeto* mutuo que trate como irrelevantes diferencias en talento y logros. (La idea de que debemos respeto al que no tiene talento, y a aquellos cuyos logros no son significativos –no le preocupó la «contribución social»– es atacada vehementemente por Nietzsche y, puesto que yo la veo como fundamentalmente correcta, es lo que hace que no comparta del todo la normal admiración hacia Nietzsche.)

A estos dos principios, parece habérsele añadido un tercero a lo largo de los siglos, a medida que la noción de «felicidad» pasó a jugar un papel más central en el pensamiento ético. Se trata de éste:

III. La felicidad o el sufrimiento de cualquiera es de igual importancia moral «prima facie».

Estos tres principios pueden ser parte de una moral totalmente secular, y han sido convertidos en ella. Por ejemplo, aunque las religiones basadas en Jerusalén consideran que el aspecto de la igualdad (el «algo en los seres humanos») mencionado en I es «ser hecho a imagen de Dios», otra elección posible, especialmente cuando el lenguaje de los «derechos» pasa a ser cada vez más utilizado acríticamente, es decir que el aspecto de la igualdad es sim-

20. En algunas versiones, debería ser insertado aquí como una cúalificación «a condición de que no pierdan este derecho cometiendo algún serio mal».

plemente que todo ser humano tiene ciertos «derechos» y que todos somos iguales en relación con estos derechos fundamentales. O, en los círculos en los cuales la noción de «felicidad» se convierte en más central que la noción de «derechos», uno puede decir que el tercer principio debería utilizarse para interpretar el primero; es decir, la igual importancia de la felicidad de toda persona *es* precisamente el aspecto en relación al cual las personas son iguales. Y, como decía, hay moralistas que piensan que toda esta idea no es más que un error; que *no* hay ningún aspecto de ninguna relevancia en relación al cual todos los seres humanos sean iguales; estos moralistas defienden varias formas de elitismo moral y rechazan la tradición moral que se deriva de los judíos y de las religiones cristianas, como hizo Nietzsche. Sin embargo, una cosa está clara: en las formulaciones tradicionales, tanto teístas como seculares, el valor de la igualdad no tiene mucho que ver con la *libertad* individual. Lo que voy a sugerir es que Kant ofreció una forma radicalmente nueva de dar contenido a la noción de igualdad, una forma que construye la libertad dentro de la igualdad. John Rawls,[21] quien también reconoce a Kant como precursor, ha realizado una interpretación del «constructivismo kantiano» que tiene una estrecha relación con la reinterpretación de Kant que yo presentaré; en efecto, argüiré que Kant era más «rawlsiano» de lo que superficialmente parece.

21. Véase de Rawls *A Theory of Justice*, Cambridge, Ma: Harvard University Press, 1971 (trad. cast. *Teoría de la justicia*, Madrid, FCE ESP, 1979) y sus Conferencias Dewey, «Kantian Constructivism in Moral Theory», que ocuparon un volumen entero de *The Journal of Philosophy*, vol. LXXVII, n. 9, septiembre 1980.

¿Cómo entiende Kant la igualdad? Según mi interpretación, el punto crucial es la presencia de cierto tipo de escepticismo en el pensamiento de Kant, un escepticismo sobre lo que se toma como más seguro en la teoría moral tradicional anterior a Kant. En un instante volveré sobre ello. Primero, permítanme decir esto: todo el mundo reconoce que la distinción central de Kant es la distinción entre autonomía y heteronomía. Pero, ¿en qué *consiste* exactamente esta distinción? La fórmula es fácil de recitar: uno es heterónomo si, por ejemplo, se ve simplemente obligado a aceptar un sistema moral, o si acepta un sistema moral sólo porque lo creen los propios padres y maestros, o si acepta un sistema moral «irreflexivamente», es decir, si no se le ocurre nunca «pensar por sí mismo». Kant es muy aficionado a la máxima «piensa por ti mismo», a la que ve como la gran máxima de la Ilustración, y frecuentemente la repite tanto en alemán como en latín. En la perspectiva de Kant, lo que la Iglesia intentó producir en la Edad Media, lo que todas las Iglesias de todos los tiempos han intentado producir y lo que los totalitarios modernos intentan producir, es gente heterónoma.

Está suficientemente claro qué es la «heteronomía». No resulta tan claro lo que es la «autonomía» o, por usar otra de las expresiones de Kant, la «autolegislación». Por supuesto, «autonomía» es lo contrario de «heteronomía». Pero lo que queremos es una caracterización positiva. ¿Qué es exactamente ser una persona autónoma, o una persona «autolegislada», en la esfera moral? Ello se parece a la libre elección, y Kant cree en la voluntad libre, pero ciertamente no es la voluntad libre en el sentido de una elección arbitraria, de un *acte gratuit*. Kant

piensa que hay una clase de razón posible en ética, y que incluso si no podemos mostrar mediante la sola razón que las premisas de un sistema ético son verdaderas, podemos mostrar que ciertos principios morales deben ser correctos *si* es que existe alguna moralidad objetiva. Éste es el famoso enfoque trascendental en ética, y aquí es donde el mismo Kant ve el paralelismo entre la segunda Crítica y la primera; pero no es aquí donde yo voy a ver el paralelismo.

La persona autónoma pregunta «Was soll ich tun?», ¿qué debo hacer? ¿cómo debo vivir?, y usa su razón para intentar imaginarse la respuesta, aunque no la use a la manera como los medievales, por ejemplo, pensaron que debían usarla. Pero hasta aquí, como vemos, es difícil apreciar cómo esto se diferencia realmente de la imagen medieval. Si una persona autónoma es aquella que es consciente de tener una libre voluntad y una razón, y aquella que usa su libre voluntad y su razón en la elección de principios éticos, entonces ¿resulta claro que esté haciendo algo diferente de lo que Tomás de Aquino le habría hecho hacer? Éste concede gran alcance a la «luz natural» y reconoce la posibilidad de llegar a alcanzar gran parte de la ética a través de la «luz natural»; así, él puede estar de acuerdo con bastantes cosas de Kant, si toda la visión de los seres humanos como autónomos se reduce a que los seres humanos tengan libre voluntad y razón. Somos, cualquier medieval lo diría, seres racionales con capacidad de libre elección; pero la noción kantiana de autonomía no es *sólo* la noción medieval de libertad más racionalidad. Únicamente viendo la diferencia específica entre las dos nociones, podremos apreciar los sentidos en los cuales la autonomía kantia-

na es una noción radical, y sólo apreciando estos sentidos, podremos ver los aspectos en los cuales Kant proporcionó un fundamento *profundo* a las nociones de igualdad que él recibió de la tradición y de Rousseau.

Para no mantenerles en suspenso, el asunto es bastante simple: de acuerdo con los medievales, como Alasdair MacIntyre[22] nos ha recordado, poseemos una capacidad para conocer la «función» humana o la «esencia» humana. Podemos también expresar su opinión diciendo que sabemos lo que es la felicidad (o la eudaemonia), entendida ésta en el «espeso» sentido aristotélico; esto es, no sólo como un sentimiento positivo, o una combinación de gratificaciones, sino como el «fin humano completo». Si esta perspectiva medieval es correcta, y conocemos en qué consiste la esencia humana, lo que es el *ergon* humano, lo que es el fin humano completo, y somos capaces de conocer todo esto usando la razón, entonces el problema de usar la racionalidad y la voluntad libre, primero para descubrir lo que uno debería hacer y después para hacerlo, es, en cierto sentido, análogo a un problema de ingeniería. El ingeniero tiene que resolver su problema, quizás con su ordenador (ya no podemos decir «con lápiz y papel»), quizás con la ayuda de algunas pruebas de laboratorio, y después tiene que aplicar la solución. De esta forma, tenemos que imaginar que se requiere que hagamos dada la función humana, dada la naturaleza de la eudaemonia, quizás llevar una vida contemplativa, o quizás vivir vidas de virtud cívica, etc. –éstos no son, naturalmente, el mismo

22. *After Virtue*, Notre Dame: University of Notre Dame Press, 1980.

ideal; el uno es griego y el otro romano–, o ser buenos cristianos o judíos o musulmanes, o cualquier otra cosa, y entonces, después de determinar esto, tenemos que hacerlo. Y se nos pueden ofrecer buenos o malos indicios, tanto para nuestro éxito como para nuestro fracaso, al imaginarnos lo que los estándares objetivos requieren de nosotros y para nuestro éxito o fracaso en vivir a la altura de ellos. Pero (según Alasdair MacIntyre), Kant entiende que esto ya no funciona. En el *Grundlegung*[23] dice explícitamente que no se puede construir una ética sobre la noción de felicidad, porque se pueden percibir demasiadas cosas en contenido de esa noción. Es sobre este *escepticismo* kantiano sobre lo que quiero llamar su atención.

Si Kant está en lo cierto, nuestra posición no es en absoluto la que Tomás de Aquino supuso para nosotros. Para ser franco, se nos pide usar la razón y la libre voluntad en una situación que, en ciertos sentidos importantes, es muy oscura, y ello debido a que la razón no nos da tal cosa como un fin humano completo que todos debamos perseguir (a menos que sea la moralidad misma, y esto no es un

23. *Grundlegung*, Zweiter Abschnitt, 418 sigs. Traducido al inglés como *The Fundamental Principles of the Metaphysic of Morals*, por Thomas K. Abbott, Nueva York; Liberal Arts Press, 1949, véase sección segunda, 35 sigs., p. ej., «But unfortunately the notion of happiness is so indefinite that although every man wishes to attain it, yet he can never say definitely and consistently what it is that he really wishes and wills». [«Pero, desafortunadamente, la noción de felicidad es tan indefinida que, a pesar de que todo hombre quiera alcanzarla, nunca puede decir de forma clara y coherente qué es lo que realmente quiere y desea» –NOTA DE LOS TRADUCTORES: cuando como ahora, lo requiera una referencia textual del autor, ofreceremos entre corchetes su traducción–].

fin que pueda determinar el *contenido* de la moralidad). El movimiento de Kant, o uno de sus movimientos –ya que pienso que aquí también hay una tensión en el pensamiento de Kant, y hay ciertamente una parte de Kant que se vuelve atrás hacia algo muy cercano a la imagen medieval– pero en el esfuerzo de Kant que estoy explicando, consiste en la idea de aceptar esta misma situación y, en lugar de disculparlo o lamentarlo, ¡decir que esto es exactamente lo que deberíamos desear! Decir que, de hecho, sería malo si hubiera un *ergon* humano revelado o una naturaleza revelada de la eudaemonia.

¿Por qué sería malo? Sería malo porque no deberíamos *querer* ser heterónomos, y la imagen medieval nos obliga a cierto tipo de heteronomía. Para mí, el pasaje más revelador de, quizás, todos los escritos morales de Kant, el pasaje en el cual mejor revela cómo la idea de esta clase de fin completo humano objetivo le repugna, está en el ensayo poscrítico titulado «La religión dentro de los límites de la mera razón». En este ensayo, Kant hace la afirmación destacable (muchos dirían que es una afirmación absurda) de que sería malo si las verdades de la religión pudieran ser deducidas mediante la razón, ya que esto produciría fanatismo.

Dense cuenta cómo Kant está usando el término «fanatismo». Está diciendo –y con una profunda intuición psicológica– que lo que hace del fanático un fanático no es que sus creencias sean necesariamente equivocadas o que sus argumentos sean incorrectos. Es posible tener creencias verdaderas, apoyadas en argumentos correctos, y todavía ser lo que Kant llama un «fanático»; es posible tener el tipo de indeseable intolerancia, intensidad, en una palabra, hostilidad a los que piensan por sí mismos,

que representaba el «fanatismo» para la Ilustración, y todavía ser perfectamente lógico. De hecho, el fanático perfectamente lógico, está diciendo Kant, es el tipo de fanático más peligroso. El fanatismo, está diciendo Kant, es indeseable por sí mismo. Que las verdades de la religión –que para Kant son las verdades más importantes– debieran ser por su propia naturaleza *problemáticas* es algo bueno, no malo. Aquí es donde la ruptura de Kant con los medievales es total.

Al mismo tiempo, Kant no está defendiendo el *fideísmo*. En el mismo ensayo, combate la idea de basar la religión en la pura «fe» como conducente *también* al «fanatismo». En efecto, Kant nos dice: «Reconozcamos que tenemos una necesidad religiosa, pero no seamos fanáticos acerca de la forma en la cual satisfacemos esta necesidad». Piensa que a través de la razón podemos convencernos del carácter *genuino* y *legítimo* de la necesidad religiosa; pero no se lamenta –por el contrario, celebra– el hecho de que la razón no pueda decirnos cómo satisfacer esa necesidad, y no piensa que deberíamos dar un «salto a la fe» hacia un fundamentalismo de algún tipo a fin de satisfacerla. Por supuesto, esto le deja (y nos deja) con un problema: ¿De dónde obtenemos un *contenido* para la religión si ni la razón ni la fe ciega pueden proporcionárnoslo? La propia respuesta de Kant es que el único contenido que estamos justificados en adoptar es nuestro derecho a *esperar*; entendiendo que la legitimación de esta esperanza es todavía un uso de la razón, aunque cualquier hipóstasis de la concepción particular de lo divino que una «iglesia» particular pueda tener como un absoluto va más allá de los límites de la razón e invita al fanatismo (o, como diríamos hoy, al

«jomeinismo»). En este sentido, la actitud de Kant es la misma que la de Lessing. «Hay una religión y muchas Iglesias», escribe Kant en algún sitio.

Como dije, encuentro terriblemente reveladores estos comentarios sobre el racionalismo y sobre el fideísmo en teología. Lo que Kant está diciendo, por expresarlo positivamente, es que tenemos que pensar por nosotros mismos sin el tipo de guía que Alasdair MacIntyre quiere restaurar para nosotros, y este hecho es, en sí mismo, el hecho más valioso sobre nuestras vidas. *Ésta* es la característica respecto de la cual somos todos iguales. Todos estamos en la misma situación, y todos tenemos el poder de pensar por nosotros mismos respecto a la pregunta ¿cómo vivir?

Si lo interpreto correctamente, la comunidad ideal de Kant es una comunidad de seres que piensan por sí mismos sin saber lo que es la «esencia humana», sin saber lo que es la «eudaemonia», y que se respetan unos a otros por hacer eso. *Éste* es el «reino de los fines» de Kant.[24]

Dense cuenta, de paso, qué lejos está Kant de la noción de sumisión a la *volunté general* de Rousseau (y esto, aunque Kant admiraba enormemente a Rousseau –el cuadro de Rousseau era el único cuadro en el estudio de Kant, y se ha llegado a mantener[25] que todo el proyecto de Kant deriva de Rousseau–). Mientras que la noción mínima de igualdad que describí antes puede ser reconciliada con varios tipos de totalitarismo, la noción de Kant no puede.

24. Compárese con las Conferencias Dewey impartidas por Rawls, citadas en la nota 3.

25. Las conferencias de Dieter Henrich, en Harvard, sobre la Deducción Trascendental de Kant son la fuente de esto.

La libertad de pensamiento es esencial, porque la característica fundamental respecto de la cual somos iguales –nuestro, por decirlo así, «aspecto de igualdad»– es precisamente nuestra necesidad de, nuestra capacidad para, un pensamiento moral libre.

Mi propuesta, entonces, es que Kant está haciendo lo que él habría llamado «antropología filosófica», o proporcionando lo que se podría llamar una *imagen moral del mundo*.[26] No está simplemente ofreciendo argumentos en favor de la tercera formulación del imperativo categórico, argumentos a favor del ordenamiento propio de los principios formales y materiales de la moralidad, etc.; está también, y principalmente, proporcionando una imagen moral del mundo que los *inspira* y sin la cual carecen de sentido. Una imagen moral, en el sentido en el cual estoy usando el término, no es una declaración de que esto o aquello sea una virtud, o de que esto o aquello es lo que se debe hacer; es, más bien, una imagen acerca de cómo nuestras virtudes e ideales se mantienen juntos unos con otros y de lo que tienen que ver con la posición en la que nos encontramos. Puede ser tan vaga como la noción de «hermandad»; realmente, millones de seres humanos han encontrado en estas metáforas imágenes morales que pueden organizar sus vidas morales –y ello a pesar del enorme problema de interpretarlas y de decidir lo que podría efectivamente significar hacerlas *operativas*.

Ahora, los filósofos morales generalmente prefieren hablar de las virtudes, o de (específicas) obligaciones, derechos, etc., más que de las imágenes

26. «Imagen moral del mundo» viene de las conferencias sobre Kant de Dieter Henrich en Harvard.

morales de mundo. Existen obvias razones para esto; a pesar de ello, creo que es un error, y que Kant está profundamente acertado. Lo que requerimos de la filosofía moral es, en primer lugar y antes que cualquier otra cosa, una imagen moral del mundo, o mejor –ya que, de nuevo también aquí, soy más pluralista que Kant– varias imágenes morales complementarias del mundo.

Pero primero, es fácil ver una similitud entre esta vertiente de Kant y el esfuerzo que me gusta de su epistemología. Para los medievales, el realismo metafísico era correcto y carecía de problemas porque se suponía que teníamos una facultad especial, la facultad de la «intuición racional», la cual nos permite un acceso directo a las cosas en sí mismas, o al menos a sus «formas». El aspecto negativo o escéptico de la primera Crítica descansa en la insistencia de Kant en que no tenemos esta clase de «intuición racional». A pesar de que Kant reconoce varios tipos de intuición, por ejemplo, las así llamadas «intuiciones puras» del geómetra, niega que tengamos la «intuición racional» medieval. La gloria de Kant, bajo mi punto de vista, consiste en decir que el mismo hecho de que no podamos separar nuestra propia contribución conceptual de aquello que está «objetivamente ahí» no es un desastre. Es, de hecho, un cierto tipo de garantía; al menos, tal como este pensamiento es reconstruido en términos contemporáneos por Strawson,[27] lo que está fundado en el argumento de Kant es que la noción de un futuro (o de un espacio-tiempo) sobre el cual

27. P. F. Strawson, *The Bounds of Sense*, Londres: Methuen, 1966 (trad. cast. *Los límites del sentido*, Revista de Occidente, 1975).

estemos *por completo* equivocados es incoherente. Si hay siquiera algo a lo que podamos llamar un «futuro», debe ocurrir que un gran número de nuestras creencias sobre él sean, de hecho, verdaderas. No quiero decir que esto sea tan cómodo como Kant pensó que era (ni tampoco Peter Strawson pretende esto), pero es algo importante, en cualquier caso, que muchos tipos de escepticismo pasen a ser, después de ser sometidos a riguroso examen, incoherentes. Quiero señalar que, de manera similar, Kant rechaza la idea de que tenemos algo análogo a la «intuición racional» medieval en relación con las cuestiones morales. Y, de nuevo, arguye que esto no es un desastre; que, al contrario, es algo bueno. La estrategia kantiana completa, en esta lectura al menos, consiste en *celebrar* la pérdida de esencias, sin volver al empirismo humeano.

El intento de la Escuela de Francfort por justificar la igualdad

Existe, sin embargo, otro intento por explicar y justificar el valor de la igualdad cuya inspiración se deriva del trabajo de Kant y que es necesario que mencione aquí; estoy pensando en el trabajo de Jürgen Habermas[28] y Karl-Otto Apel.[29] Este trabajo es demasiado complejo para discutirlo en detalle, pero las ideas principales son bien conocidas. Habermas

28. Jürgen Habermas, *Knowledge and Human Interests*, traducido al inglés por Jeremy J. Shapiro, Boston: Beacon Press, 1971 (trad. cast. *Conocimiento e interés*, Madrid, Taurus, 1986).

29. Karl-Otto Apel, *Charles S. Peirce, from Pragmatism to Pragmaticism*, traducido al inglés por John Michael Krois, Amherst: University of Massachusetts Press, 1981.

y Apel pretenden –y yo estoy de acuerdo– que la noción de una afirmación justificada o con garantías involucra una referencia implícita a una comunidad. Más aún, ellos creen que abordar un discurso que pretende ser racional nos compromete implícitamente con ciertas exigencias: exigencias como, por ejemplo, la verdad, la sinceridad, la justificación, etc., en relación con lo que se dice. Explican que la idea de una afirmación completamente *justificada* es la idea de una afirmación que puede resistir los controles y la crítica. (Ésta es una noción de afirmación justificada «atemporal» o absolutizada, no la noción de una afirmación que está «justificada de acuerdo con la evidencia presente», o «justificada para alguien en tal o cual posición».) Incluso cuando hacemos una afirmación que ordinariamente no puede ponerse en duda, por ejemplo, «estoy levantando mi mano», se pueden pensar casos poco probables en los cuales la afirmación pueda ponerse en duda con éxito. Hubo una famosa ocasión en la cual G. E. Moore ofreció como ejemplo de certeza «sé que hay una ventana en esta habitación» y en la cual estaba, de hecho, equivocado (Moore estaba dando una conferencia en un salón de la Universidad de Michigan que tenía cortinas, pero ninguna ventana detrás de las cortinas). La idea de una afirmación cuya garantía completa y final sea totalmente accesible al propio hablante *al margen de lo que ocurra* –o de un hablante que ni necesite ni pueda beneficiarse de los datos de otros– es justamente la vieja noción de conocimiento privado e incorregible. La reflexión interesante es que si Peirce y Wittgenstein están en lo cierto (y Apel se refiere mucho a ambos), y el conocimiento privado y el conocimiento incorregible son ideas vacías y falaces,

entonces no puede haber nada como una afirmación que sea verdadera (o, de todos modos, una afirmación que sea verdadera humanamente hablando, es decir, capaz de resistir todos los intentos posibles de falsearla) a menos que exista la posibilidad de una comunidad de comprobadores o, en cualquier caso, de críticos.

El resultado es que *si soy una persona racional, en el sentido de tener el propósito de realizar afirmaciones que sean verdaderas humanamente hablando, es decir, que puedan soportar la crítica racional ahora y en el futuro, entonces estoy comprometido con la idea de una posible* comunidad *de investigadores*. (De hecho, con la idea de una comunidad posible de tamaño potencialmente infinito, ya que puede que no haya tal cosa como un investigador último si se permite que pueda reabrirse cualquier investigación.)

Hasta aquí, puede que esto parezca no tener nada que ver con la ética, aunque tenga mucho que ver con las concepciones pragmatistas de la razonabilidad. Pero, ¿qué *clase* de comunidad debe ser esta comunidad ideal? Si digo algo proponiéndome la verdad y la garantía, entonces estoy, de acuerdo con Apel y Habermas, *reconociendo la autoridad* (al menos, en principio) *de una comunidad que sigue funcionando, una comunidad que debe tener cierta estructura*. Una comunidad que sea competente para determinar la verdad y la falsedad debe ser tal que, por ejemplo, cualquiera en esa comunidad pueda *criticar* lo que se exponga sabiendo que su crítica será escuchada; si algunas críticas simplemente no son oídas, entonces la posibilidad de un tipo irracional de «protección de la creencia» asoma su fea cabeza; estamos volviendo a lo que Peirce llamó el «método de la autoridad». Nuevamente,

no sólo debe ser posible para cualquier miembro de la comunidad hacer una pregunta o elevar una crítica; debe ser también posible para cualquier miembro de la comunidad de investigadores ideales avanzar una hipótesis sabiendo que será escuchada. Debe ser una comunidad, en una palabra, que respete los principios de la libertad e igualdad intelectuales.

Todavía se puede decir que es posible reconocer el valor de tener comunidades que obedecen a una ética de la igualdad y la libertad intelectual mientras que, al mismo tiempo, no creen que esos principios deberían ser universalizados. Se puede mantener esclavos, como hicieron los antiguos griegos, mientras que, al mismo tiempo, se discute con gran libertad intelectual con un colega dueño de esclavos si la esclavitud está o no justificada. Sin embargo, como Habermas ha insistido, después de todo hay algo de «falsa conciencia» en la noción marxista: la investigación de Aristóteles sobre la ética de la esclavitud no era realmente «libre»; la conclusión era inevitable.

De nuevo, se puede pretender que la necesidad de que las comunidades de investigadores respeten los principios de igualdad y libertad intelectual no implica realmente que las sociedades tiránicas sean, en ningún sentido, «irracionales». Después de todo, un tirano puede querer tener una comunidad de científicos, e incluso puede permitirles tener cierta libertad de expresión, en tanto su libertad de expresión se restrinja al laboratorio. Pero el tirano no sentirá que esto le compromete con el valor de la libertad de expresión en la sociedad civil. Apel y Habermas replicarían señalando que aquél no está permitiendo que los científicos tengan libertad de

expresión sobre cuestiones políticas y sociales, y que en esa medida no está permitiendo *completa* libertad de investigación. Si el tirano pretende que sus propias creencias políticas sean racionales a la vez que las protege contra la crítica racional, entonces se involucra en una especie de contradicción.

He dicho suficiente para indicar por qué pienso que la posición de Apel y Habermas es extremadamente interesante, y merecedora por parte de los filósofos analíticos de más atención de la que normalmente recibe. Pero la conclusión que quiero sugerir, a partir de lo que pienso que son las auténticas intuiciones de Apel y Habermas, no es que se pueda derivar una ética universal a partir de la idea de razonabilidad, o a partir de la imagen (kantiana como también peirciana) de la verdad como aceptabilidad racional idealizada, sino una diferente, más débil pero aún importante, conclusión. Habermas está en lo cierto: existe una tensión real entre desear que las propias creencias sean completamente racionales y tolerar una sociedad que restrinja, de manera fundamental, la libertad intelectual. También existe otra tensión involucrada en pretender tener una sociedad que no restringe para nada la libertad individual a la vez que se justifica cualquier forma de opresión.

Un problema, sin embargo, en relación al intento de derivar una ética universal a partir de estas consideraciones consiste en que el oponente puede *no* pretender que sus creencias éticas sean racionales. Realmente, si él es un «no-cognitivista», puede negar que *cualesquiera* creencias éticas sean o puedan estar justificadas racionalmente. Esto, en sí mismo, no preocuparía a un filósofo que estuviera argumentando según la tradición kantiana, ya que

aquél podría (como el mismo Kant) contentarse con mostrar que ciertos valores están racionalmente justificados *si* es que existen en verdad valores racionalmente justificados. Pero aparece un segundo problema cuando se argumenta contra un oponente que es un elitista: el oponente puede insistir en que las *mismas* sociedades estructuradas democráticamente producen «falsa conciencia». Puede decir que el hecho de que ciertos grados de igualitarismo y libertad intelectual sean necesarios en una comunidad de científicos físicos no muestra que sean necesarios o deseables en una comunidad cuyo objetivo sea conseguir el superhombre; un nietzscheano insistirá en que sólo a través de cierto tipo de sociedad *no igualitarista* puede despuntar la verdad sobre la existencia humana.

Estas objeciones no muestran que los intentos de Apel y Habermas de derivar los valores de la igualdad y la libertad intelectual a partir de la racionalidad no tengan ningún valor. Sus argumentos *muestran* que nuestros valores de la igualdad, la libertad intelectual y la racionalidad están profundamente interconectados. Sin embargo, la moral que quiero extraer de ello es más aristotélica que kantiana: tenemos una idea rica y polifacética del bien, y las partes de esta idea son interdependientes.

Aunque, como dije, el realismo interno tenga una conexión familiar con el kantismo, la forma en la que divergiré en mayor medida de Kant radica en no intentar (ni incluso pretender intentar) derivar la ética de la sola razón. Este proyecto fue parte del sueño de una «ética universal» del siglo XVIII –un sueño que no estaba siempre separado, como Sidney Morgenbesser ha señalado en conversación, del sueño de una «forma de vida universal»–. A pesar de

que la relación de lo razonable con lo ético sea parte de mi historia, creo que hay otra parte de la historia que merece atención –otra parte que, paradójicamente, también proviene de Kant.

Democracia sin una imagen moral

Una manera de ver la importancia de una antropología filosófica, de una imagen moral del mundo, es ver qué resulta cuando intentamos justificar nuestras instituciones sociales fundamentales –especialmente la institución llamada «democracia»– sin apelar a una imagen de la naturaleza humana. Quizás la defensa más común de la democracia sea la de Winston Churchill. Estoy refiriéndome al famoso chiste «La democracia es el peor de todos los sistemas –exceptuando, naturalmente, todos los sistemas que ya han sido probados». Winston Churchill estaba ofreciendo un argumento real a favor de la democracia –todo lo demás que conocemos funciona *peor*–. Pero esta defensa, a pesar de ser fuerte e importante, sería vulnerable si apareciera una tecnocracia auténtica con medios para condicionar a sus ciudadanos a aceptar sus valores (y a sentirse «felices»). Recientemente, he leído por primera vez la profunda novela de Huxley *Un mundo feliz*, y encontré que era un libro muy importante filosóficamente. Así que intenté discutirla con mis amigos, todos los cuales la habían leído. Por desgracia, ocurrió que la habían leído en la educación secundaria, incluso en la educación primaria, y «no la recordaban muy bien». En consecuencia, todavía no he sido capaz de discutirla con nadie que recuerde el leerla y el reaccionar a ella como un adul-

to. Es una gran lástima que este libro haya sido relegado a los años de la educación secundaria (donde puede ser presentado como una crítica de «ellos» –los totalitarios– y no como una crítica de *nosotros*), ya que el diálogo casi al final del libro con Mustafa Mond, el Controlador del Mundo, aunque no sea, tal vez, una gran pieza de literatura como el famoso diálogo con el Gran Inquisidor, es una contemplación de ficción bellamente conseguida y de profundas consecuencias. Recordarán (espero) que en *Un mundo feliz* la gente es, en cierto sentido, «feliz». Ellos ciertamente disfrutan de un alto «tono hedónico» y tienen una enorme cantidad de «diversión». Sin hacer trampas, es difícil ver cómo el utilitarismo clásico basado-en-el-placer podría negar que, si ellos no han *maximizado* realmente la «felicidad», han conseguido ciertamente mucha más felicidad en *ese* sentido de la que tenía la sociedad que la Distopía de Huxley reemplazaba. El argumento de que la democracia funciona mejor que las alternativas conocidas funciona bien mientras que las alternativas que conocemos conduzcan a guerras, pobreza y gran sufrimiento, etc. El precio, en la novela, es una deliberada *infantilización* de la raza humana. La gente está condicionada a permanecer emocionalmente adolescente durante toda su vida, y a disfrutar, mientras vivan, exclusivamente de los placeres de la adolescencia. Si se trajera a sus ciudadanos al momento presente, y se les permitiera ver nuestra actual sociedad democrática con sus nerviosismo y tensiones, sus conflictos emocionales, sus conflictos políticos, etc., de acuerdo con la premisa de la novela la gran mayoría de ellos querría volver a sus placeres infantiles (y a su droga «soma»).

He dicho que el utilitarismo clásico no tiene realmente una respuesta no tramposa a la pregunta «¿qué hay de malo con el mundo feliz de Huxley?» (adoptando las premisas de la novela). Se puede pensar que, sin embargo, el utilitarismo moderno la tiene. El utilitarismo moderno (lo llamaré el utilitarismo de la «función-de-preferencia») está basado en la idea de dar a la gente lo que ellos *prefieran* (de aquí las referencias frecuentes a la noción de una «función de preferencia» de los economistas). Lo que es incorrecto en la Distopía de Huxley –podría decir un utilitarista moderno de la función-de-preferencia– es que la gente (esto es, la gente *ahora*) no *preferiría* vivir en el mundo feliz, y el utilitarista de la función-de-preferencias insiste en que la felicidad ha de ser determinada por las funciones de preferencia de la gente *tal como es ahora* –no como sería después de un adecuado condicionamiento–. Pero lo que la novela de Huxley señala es algo que a menudo se olvida en estas discusiones. Lo que se olvida es que la función-de-preferencia utilitarista es simplemente una manera de «introducir de contrabando» en el utilitarismo otro valor distinto de la felicidad; es decir, el valor de la *libre elección*, sin proveer a la imagen de ninguno de los recursos para defender ese valor que podrían derivarse de una imagen «más densa» de la naturaleza humana.

Lo que tengo en mente es esto: si Bernard, en su diálogo con el Controlador del Mundo, hubiera dicho que la sociedad es mala porque no es la sociedad que la gente en la época precedente habría escogido si hubieran tenido posibilidad de elegir (recuérdese que la época precedente era una época de guerra química y biológica de la escala de un holocausto), Mustafa Mond simplemente se habría reí-

do. Después de todo, se trata de la «mejor sociedad» de acuerdo con las funciones de preferencia de la gente que vive en ella, aunque esas funciones de preferencia sean el producto de un condicionamiento. Sin duda, ese condicionamiento es algo sólido, pero no es toda la historia. Los placeres infantiles que disfrutan, las drogas, el sexo, las «emociones», etc., son placeres reales; los sufrimientos que proceden del compromiso emocional, de la lealtad y la amistad, de los compromisos religiosos y políticos, etc., son sufrimientos reales. Esa gente *tal como es* en *su* mundo prefiere genuinamente su mundo. Es el mundo que «maximiza la felicidad» de acuerdo con los estándares proporcionados por las funciones-de-preferencia que ellos actualmente tienen. ¿Por qué deberíamos convertir en un fetiche la satisfacción de las funciones de preferencia que la gente tiene sin condicionamiento, en oposición a hacerlo felices aunque hacerlos felices requiera *cambiar* esas funciones de preferencia? Mustafa Mond argumenta que se han intentado experimentos controlados, experimentos empleando la parte más inteligente de la población y ofreciéndoles la oportunidad de formar una sociedad libre. Esos experimentos acabaron una y otra vez, nos dice, en guerra y destrucción. Lo importante es que Mustafa Mond *tiene* una imagen del mundo, a pesar de que la encontremos repulsiva –una imagen naturalista y reduccionista– y tiene una fuerte defensa, a la luz de su imagen, respecto a que está claro que su sociedad hace a la gente *feliz*.

He dicho hace un momento que la función-de-preferencia utilitarista realmente es una forma de introducir de contrabando un segundo valor en el utilitarismo. Mi hipotético diálogo entre un utilita-

rista de la función-de-preferencia y Mustafa Mond (que no está muy lejano del diálogo real del libro) fue diseñado para aclarar esto. En efecto, el utilitarista de la función-de-preferencia insiste en que tenemos que hacer feliz a la gente, pero debemos hacer esto sin dañar su libertad para decidir por sí mismos lo que *es* su felicidad. Esto no es un valor (la «felicidad»), sino dos; de hecho, son tres ya que uno de los principios que he enumerado antes en mi descripción de la «mínima noción de igualdad basada-en-Jerusalén», esto es, el principio de que *la felicidad de cualquiera (o la carencia de ella) es de igual importancia «prima facie»*, es asumido por todas las formas de utilitarismo.[30] El utilitarismo es un intento por hacer que una serie de ideas que tienen profundas y complejas raíces en nuestra cultura, los valores de igualdad, los valores liberales de elección, y los valores de fraternidad y felicidad, parezcan simples y no arbitrarios. Como he dicho antes, cuanto más se intente hacer esto sin examinar esos valores por separado, y examinando las imágenes del mundo en que puedan basarse esos valores, peor nos va a ir; ya que si todos esos valores descansan en una noción general no analizada de que «la gente debe ser feliz», entonces, como digo, el desafío de Mustafa Mond, el desafío de alguien que pregunta «¿por qué uno que está pensando *científicamente* sobre la felicidad debería suponer que

30. Sin embargo, este igualitarismo no fue siempre una parte de la doctrina utilitarista. «Cada cual cuenta por uno, ni por más ni por menos» fue una adición de Bentham: Hutcheson, quien anticipó a Bentham con un cálculo hedónico, trabajaba con la idea de que «la dignidad, o la importancia moral de las personas, puede compensar a los números» –p. ej., alguna gente que fuera más importante podía contar por más.

cualquier Tom, Dick o Harry es un buen juez de su propia felicidad?» se convertirá rápidamente en algo sin respuesta. No sé, de hecho, qué es más terrible: la Distopía descrita por Huxley, que la describe con disgusto, o la utopía descrita por Skinner en *Walden dos*, que la describe con una apetencia positiva.

Quiero puntualizar que la defensa de los valores liberales, especialmente la defensa de la libertad individual, en la cual Kant estaba muy interesado, no puede ser confiada al utilitarismo si la única forma como el utilitarismo puede defenderlos es pasándoles de contrabando en su definición de «felicidad», lo cual es, después de todo, un camino intelectualmente sin defensa. El utilitarismo es una visión inadecuada justamente por las razones que Kant dio: la inutilidad de la noción de «felicidad», y la insensibilidad de las éticas basadas-en-la-felicidad para sacar conclusiones acerca de los *medios*.

La de Kant no es, por supuesto, la única imagen moral alternativa que tenemos. Las religiones basadas-en-Jerusalén tenían una imagen que, aunque no incluía aún los valores liberales del pensamiento libre y crítico, enfatizaban la igualdad y también la fraternidad, como en la metáfora de toda la raza humana como «una familia», de todas las mujeres y hombres como hermanos. Y yo mismo encuentro la filosofía moral de Kant defectuosa en dos aspectos. Kant no quiere complementar las imágenes morales anteriores, sino reemplazarlas completamente desde un punto de vista moral monista. Necesitamos una visión más plural que la que el siglo XVIII podía prever. Puedo también simpatizar con aquellos que piensan que la sustitución de una noción de fraternidad que enfatiza la idea de tener *sentimien-*

tos fraternales hacia los demás por una noción de fraternidad que enfatiza hacer lo que se *debe* hacia los demás entraña una pérdida real. Así, estoy en desacuerdo con el rigorismo de Kant *y* pienso que necesitamos una imagen moral del mundo más polifacética. No sólo tenemos una imagen basada-en-Jerusalén fundada en la fraternidad –que pasa a secularizarse en la medida en que «libertad, igualdad, fraternidad» se convierte en *el* gran eslogan de la Revolución francesa–; tenemos aún otra imagen procedente del pensamiento romano, y revivida hoy por algunos pensadores liberales e incluso de izquierdas, la imagen del «republicanismo cívico» romano, que enfatiza la idea del *homo sapiens* como una criatura cuya dignidad e identidad se deriva de su pertenencia a una *polis*, y cuyo más elevado derecho no es la «contemplación», sino el ejercicio de las virtudes cívicas. Tenemos imágenes morales muy diferentes en nuestra tradición, y hay también imágenes morales en otras tradiciones que merecen atención. Sin embargo, me parece que la imagen que Mustafa Mond propone en *Un mundo feliz*, o la imagen que Skinner propone con toda seriedad en *Walden dos* –la imagen de la felicidad humana como un mero problema *técnico*– empezará siempre por afirmar su «validez científica».

Como he señalado, describir un ideal como una imagen moral es decir que es mucho más que una caracterización de algún rasgo o de algún modo de comportamiento virtuoso. Ciertamente, la imagen moral kantiana incluye la pretensión de que pensar por uno mismo sobre cómo vivir (*für sich selbst denken!*) es una virtud, pero contiene también otras muchas ideas. Contiene la idea de que esta virtud no es solamente una virtud, sino que nuestra capa-

cidad para ejercitar esta virtud es la capacidad moral más significativa que tenemos; incluye la afirmación de que un ser humano que ha elegido no pensar por sí mismo sobre cómo vivir, o que ha sido forzado o «condicionado» a ser incapaz de pensar por sí mismo sobre cómo vivir, fracasa en desarrollar una vida completamente humana. Contiene también la visión de una comunidad de individuos que se respetan unos a otros por esta capacidad, e ideas sobre cómo tal comunidad debería organizarse y sobre la posición moral que ocupan los miembros de tal comunidad, sobre lo que pueden y no pueden conocer. En una palabra, es una visión que incluye y organiza un sistema complejo de valores. Lo mismo es verdad del sistema de creencias que constituye la tradición del republicanismo cívico que he mencionado hace un momento. E, incluso la metáfora de la «familia» humana, de la hermandad universal, aun siendo tan vaga, ha servido para organizar el pensamiento moral y para dar significado a la moralidad durante cientos de generaciones de seres humanos.

Sin embargo, si esto es correcto, entonces nuestro problema puede parecer más difícil, no más fácil. Los «valores» individuales, los derechos, las virtudes, etc., son suficientemente difíciles de justificar; ¿cómo se podría esperar justificar una visión moral completa, una «antropología filosófica»? Tanto la complejidad como la elasticidad, o la posibilidad de diferentes interpretaciones, de los objetivos que son organizados por cualquier imagen moral se convierten en una fuente de dificultad. Pero una discusión de *este* tema se pospondrá hasta la próxima y última conferencia.

CONFERENCIA IV

LA RAZONABILIDAD COMO HECHO Y COMO VALOR

En este punto, no me sorprendería si muchos de mis oyentes se sintieran inclinados a decir algo como esto: «Las imágenes morales que ha descrito en la última conferencia son espléndidas, maravillosas, pero ¡atención! Cualquier filósofo serio preguntará cómo podemos *justificar* algo de esto».

En un libro que publiqué hace algunos años,[31] defendí la idea de que algo puede ser a la vez un hecho y un valor –dije que es un hecho, por ejemplo, que Yeats fue un gran poeta, y un hecho que los nazis fueron malvados–. Y, también entonces, fue la cuestión de la justificación la que preocupaba a la gente. Muy a menudo, la gente expresaba su preocupación preguntando: «Pero, ¿no tiene que admitir que hay mucho más acuerdo en los resultados científicos que en los valores éticos? ¿No muestra esto que hay una clase de objetividad que tienen los resultados científicos y de la que carecen los valores éticos?».

Un argumento está implícito en la pregunta. Podemos llamarlo el *argumento a partir de la no conflictividad*. La idea es que la marca de *status* cognitivo está, de alguna forma, en la posibilidad de convertirse en conocimiento «público», es decir, de llegar a ser algo no controvertido.

31. *Reason, Truth and History*, Nueva York: Cambridge University Press, 1981.

No quiero sugerir que alguien realmente piense que sólo lo que no es controvertido es realmente conocimiento. La idea, más bien, es que los «hechos» pueden ser demostrados «científicamente». Si existe controversia sobre una cuestión fáctica es porque aún no hemos realizado suficientes experimentos, o acumulado suficientes datos. Lo que sea un hecho puede «en principio» ser establecido de manera que exija el asentimiento de todas las «personas racionales», donde a menudo se considera que esto significa que todas las personas *cultas*, o todas las personas *inteligentes*. Así, Weber[32] argumentó que los juicios de valor sólo pueden ser un asunto de «fe»; y pensó que esto quedaba suficientemente demostrado al señalar que existían algunos valores occidentales acerca de cuya corrección no podía convencer a un «mandarín chino». El tema con el que quiero comenzar hoy es este supuesto carácter «público» de los hechos, la idea de que los hechos pueden «en principio» establecerse más allá de la controversia.

No está claro en absoluto que esta idea sea correcta, ni siquiera para las «ciencias duras». La ciencia ha cambiado de opinión de forma sorprendente acerca de la edad del universo, y puede hacerlo de nuevo. Si determinar algo más allá de la controversia es hacerlo *de una vez por todas*, en oposición a establecerlo meramente de manera que se convierta en conocimiento aceptado por una época, entonces no resulta nada claro cuánta ciencia fundamental se ha creado, o incluso se creará, «más allá de la controversia». La conclusión es algo común para los filósofos de la ciencia a partir de dis-

32. Véase *The Methodology of Social Sciences*, Nueva York: The Free Press (MacMillan), 1949.

cusiones relativas a la «convergencia», y no pretendo proseguirla aquí.

Una razón por la cual la pregunta de mi interlocutor no puede ser rechazada señalando lo «controvertido» que es el tema de la convergencia en la ciencia teórica radica en que el que pregunta probablemente dirá que hay convergencia al menos respecto de los resultados *observacionales*. Aquí, por supuesto, también hay problemas. En otro tiempo, la «simultaneidad» parecía ser un término observacional (en predicciones tales como «si haces X, el sistema hará simultáneamente Y y Z»). Pero, desde la aceptación de la Teoría de la Relatividad Especial, somos conscientes de un componente fuertemente «teórico» en la noción de simultaneidad. Puede hacerse una indicación similar acerca de otros términos «observacionales».[33] Pero mi interlocutor no me estaría haciendo esta pregunta si creyera la historia de Thomas Kuhn[34] acerca de la «inconmensurabilidad» de incluso el vocabulario observacional de las ciencias. En consecuencia, dirá justamente que incluso si nuestra *teoría* de la simultaneidad, o de cualquier otra cosa, ha cambiado, aún hay *algo* invariante acerca del tipo de predicción que acabo de mencionar.

Permítaseme, mejor, *conceder* justamente a mi

33. Por ejemplo, la *forma* de un objeto depende, en la Teoría Especial de la Relatividad, de la velocidad del observador en relación al objeto. «Cuadrado» es así, estrictamente hablando, un término *dependiente del marco* en la física relativista. Esto ilustra el hecho de que ningún término en el «vocabulario observacional» ordinario del físico sea inmune a la posibilidad de recibir clases extremadamente complejas de carga teórica.

34. Véase *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press, 1981.

interlocutor su idea de que conseguimos convergencia, consenso, libertad de controversia, en las ciencias duras, al menos «en principio». La pregunta que merece un análisis más profundo es «¿Qué se sigue?». La presencia en otras áreas, áreas que no cuentan como «auténtica ciencia», de disputas que no pueden resolverse, ¿muestra realmente que se están haciendo afirmaciones «subjetivas» o «sin significado cognitivo»? Y (ya que he estado enseñando recientemente a William James), ¿cuáles serían las consecuencias para nuestras vidas si respondiéramos *seriamente* «sí» a esta pregunta? ¿Cuál es el «valor en efectivo» de creer que sólo lo que puede ser establecido más allá de la controversia tiene algo que ver con la «cognición», el conocimiento, la comprensión?

No es poco. Considérese, por un momento, nuestro conocimiento *histórico*. Tuve ocasión de leer recientemente un trabajo de un gran historiador,[35] quien construyó una potente defensa de la concepción según la cual la civilización «medieval» no comenzó realmente con la caída de Roma. Bajo el reinado merovingio, arguye este historiador, la civilización que prevaleció era aún reconociblemente «romana». Los cambios drásticos que crearon lo que todavía hoy nos.representamos como civilización «medieval» –la desaparición del latín, excepto entre monjes muy estudiosos, la ascensión de los príncipes germanos del norte, la ruptura completa con Bizancio, todo el orden feudal– llegó como con-

35. Henri Pirenne, *Mohammed and Charlemagne* (traducción al inglés de su *Mahomet et Charlemagne*), Totawa, Nj: Barnes and Noble, 1980 (trad. cast. *Mahoma y Carlomagno*, Madrid, Alianza, 1985).

secuencia del surgimiento del Islam, y de la conquista islámica de Egipto, el África romana y España. Éste es un típico ejemplo de reconstrucción histórica. Todos sabemos que tales reconstrucciones son, por regla general, extremadamente controvertidas. Imaginemos que esta reconstrucción particular continúa siéndolo. ¿Qué tenemos que concluir?

Si no tomamos en serio la idea de que la conflictividad indica falta de *status* «cognitivo», entonces estaremos obligados a creer que este historiador no ha hecho ninguna afirmación con «significado cognitivo» –¡una conclusión que parece absurda a primera vista!

Mi interlocutor puede responder diciendo que el principio de que lo que es un hecho puede ser establecido más allá de la controversia sólo se pretende aplicar a las afirmaciones fundamentales de existencia. Si las afirmaciones fundamentales de existencia presupuestas por una investigación o por una disciplina completa han sido verificadas «públicamente», entonces las hipótesis individuales formuladas en términos de las entidades y propiedades cuya existencia ha sido así establecida no necesitan ser puestas a prueba individualmente para contar como cognitivamente significativas. (Los empiristas lógicos hicieron un movimiento de esta clase cuando cambiaron desde la «verificabilidad» a la «expresabilidad en un lenguaje empirista» como «criterio de significación cognitiva».)[36] La existen-

36. Véase el clásico trabajo de C. G. Hempel, «Problems and Changes in the Empiricist Criterion of Meaning», en *Revue International de Philosophie*, 4: 41-63 (1950), reimpreso en L. Linsky (comp.), *Semantics and the Philosophy of Language*, págs. 163-185, Urbana: University of Illinois Press, 1952.

cia de la Edad Media, de los merovingios, de Pippin y Carlomagno y el Papa y Bizancio, etc., no es algo controvertido, podría señalar, y esto justifica nuestra consideración de las hipótesis individuales sobre ellas como significativas cognitivamente.

Sin embargo, hay dos problemas (al menos) con esta respuesta. En primer lugar, la hipótesis en cuestión es una hipótesis sobre las causas de todo un marco social, una hipótesis sobre «por qué ocurrió lo que ocurrió» o, si no, sobre lo que no habría ocurrido si el surgimiento del Islam no hubiera tenido lugar. ¿Ha sido «científicamente» establecida la existencia misma de explicaciones causales globales, o de hechos sobre lo que habría ocurrido si un cierto evento histórico irrepetible no hubiera ocurrido? Es difícil ver cómo. Realmente, hay filósofos para los que ninguna de las afirmaciones de este tipo no tendría significado cognitivo. En segundo lugar, como señalé en la primera conferencia, la lista de objetos cuya «existencia real» no ha sido establecida por la «ciencia dura» (de acuerdo con destacados filósofos) es realmente impresionante –¡incluso los cubitos de hielo no están en ella, de acuerdo con Sellars!–. De hecho, lo que la ciencia dice sobre el *comportamiento* de los cubitos de hielo es mucho más claro que si la ciencia dice que «realmente existen».

Una respuesta diferente fue la ofrecida hace mucho tiempo por Ernest Nagel[37] y también por Rei-

37. Por ejemplo, Nagel escribió que «There appears to be no good reason for claiming that the general pattern of explanations in historical inquiry ... differs from those encountered in the generalizing and the natural science» [«No parece existir una buena razón para afirmar que el patrón general de las explicaciones en la investigación histórica ... difiere de aquellos

chenbach[38] y otros empiristas de su escuela. De acuerdo con estos filósofos, las inferencias hechas por los historiadores son simplemente inferencias inductivas del mismo tipo de las que vemos en física. Si esas inferencias nos conducen a grados de confirmación muy altos (según estos pensadores),

---

encontrados en las generalizaciones y en la ciencia natural»] («The Logic of Historical Analysis», en H. Feigl y M. Brodbeck (comps.), *Readings in the Philosophy of Science*, Nueva York: Appleton-Century-Crofts, 1953, 688-711). Inmediatamente después de hacer esta afirmación, Nagel continúa diciendo que las «explanatory premises in history» [premisas explicativas en historia] incluyen leyes «as well as many explicitly (although incompletely) formulated statements of initial conditions» [«así como muchas declaraciones explícitamente formuladas (aunque de forma incompleta) acerca de las condiciones iniciales»]. Nótese que «law» [«ley»] e «initial conditions» [«condiciones iniciales»] pertenecen a la jerga de los físicos, y no a ninguna manera de hablar histórica.

38. Por ejemplo, Reichenbach escribe que «The argument that sociological happenings are unique and do not repeat themselves breaks down because the same is true for physical happenings. The weather of one day is never the same as that of another day. The condition of one piece of wood is never the same as that of any other piece of wood. The scientist overcomes these dificulties by incorporating the individual cases into a class and looking for laws that control the unique conditions at least in a majority of cases. Why should the social scientist be unable to do the same thing?». [El argumento de que los acontecimientos sociológicos son únicos y no se repiten muestra ser falso porque lo mismo es verdadero de los acontecimientos físicos. El tiempo que hace un día nunca es el mismo que el de otro día. El estado de un trozo de madera nunca es el mismo que el de otro trozo de madera. El científico vence estas dificultades subsumiendo los casos individuales a una clase y buscando leyes que controlen las condiciones únicas al menos en una mayoría de casos. ¿Por qué tendría que ser incapaz el científico social de hacer lo mismo?] *The Rise of Scientific Philosophy*, Berkeley: University of California Press, 1951, 309-310.

es porque la ciencia social aún está «inmadura»; no hemos descubierto todavía las leyes de la conducta social. Cuando tengamos éxito al formularlas, entonces seremos capaces de determinar de manera más satisfactoria cuáles de las interpretaciones de los eventos sociales disponibles en el presente son aceptables y cuáles no.

Hoy día, sin embargo, la misma creencia de que *hay* «leyes de la conducta social» de un tipo preciso –y la creencia de que el destino de la ciencia social es parecerse a la física– se ha convertido en una opinión minoritaria entre los científicos sociales y también entre los filósofos y metodólogos de las ciencias sociales. Ciertamente, ésta es una creencia empírica; si es falsa, entonces las teorías históricas como la que tomé como ejemplo no pueden, en este sentido, ser verificadas más allá de una controversia significativa.

Reichenbach, pienso, no habría estado tan preocupado como Nagel por esta posibilidad. Mis recuerdos de él, en la época en la que yo era estudiante suyo, me conducen a pensar que él podría haber dicho algo como lo siguiente: «A veces, tenemos que aceptar teorías que sólo están débilmente confirmadas, cuando todas las alternativas son incluso más improbables. El hecho de que las teorías históricas no tenga un grado alto de probabilidad, y quizás nunca lo tengan, no quiere decir que la naturaleza de las inducciones que usamos para medir esta probabilidad sea de alguna manera especial».

Si esta concepción es correcta, entonces lo que hay de especial en la «ciencia» no es en absoluto la posibilidad de establecer la verdad de las hipótesis de forma más o menos conclusiva (más allá de la controversia). Esta perspectiva concede una de las

cosas a favor de las que quiero argumentar: que una hipótesis o una afirmación puede estar garantizada, puede ser razonable creer en ella, en un sentido objetivo de las palabras «garantía» y «razonable», aunque no podamos especificar un experimento (o datos) tal que si lo realizáramos (o si lo obtuviéramos) fuéramos capaces de confirmar o no confirmar la hipótesis en una medida que obligara a la conformidad a toda la gente culta e imparcial. De acuerdo con la concepción de Reichenbach,[39] lo que es esencial a la ciencia es su uso de la «inducción», y en modo alguno la posibilidad de conseguir que la mujer de la calle (o la mujer culta de la calle) crea sus resultados.

Mi interlocutor podría, por supuesto, decir que sería posible contar la controversia sobre las causas de los hechos característicos de la civilización «medieval» *produciendo un gran número de mundos posibles justamente como el nuestro antes del nacimiento de Mahoma, pero en los cuales Mahoma no hubiera nacido, y ver lo que ocurre*, pero esto sería ciencia ficción más que metodología. O él podría decir que «puede aparecer alguna evidencia que no tenemos ahora en cuenta pero que decidirá el resultado –podemos incluso descubrir aquellas leyes de la ciencia social esperadas por Ernest Nagel–». Pero es totalmente distinto decir que un tema es «decidible en principio», en el sentido de que se puedan describir experimentos –experimentos que sea humanamente posible realizar– cuyo resultado confirmará o no confirmará la hipótesis en un alto grado, o decir «puede aparecer algo». Después de

39. Véase, por ejemplo, *Experience and Prediction*, Chicago: University of Chicago Press, 1938.

todo, hay pensadores[40] que mantienen que la mayor parte del desacuerdo ético es un efecto colateral no reconocido del desacuerdo fáctico: si la gente estuviera de acuerdo en la teoría psicológica y social, en el sentido más amplio (concepciones de la naturaleza humana y de la sociedad), entonces, estos pensadores afirman, estarían de acuerdo en la mayoría de los temas éticos en discusión. No es lógicamente imposible que esta perspectiva sea correcta; en consecuencia, «en principio» puede aparecer algo que acabe con el desacuerdo ético. Ésta no es la pregunta que la persona que pregunta si podría acabarse con el desacuerdo ético está haciendo: está preguntando si podemos especificar *ahora* algún descubrimiento que acabe con las controversias éticas.

Algunas veces, de hecho, se está pidiendo algo incluso más fuerte. Robert Nozick cuenta que se le preguntó si se podría dar un argumento para mostrar que Hitler era una mala persona, *que convenciera a Hitler mismo*. La única respuesta a su pretensión –la pretensión de que lo que sea un hecho debe poder ser probado a toda persona «inteligente»– consiste en señalar que probablemente ninguna afirmación excepto el principio de no contradicción tenga esta propiedad.[41]

Por último, mi interlocutor podría ceder a la tentación y negar que las hipótesis históricas (como

40. Richard Boyd ha argüido recientemente a favor de este punto de vista en conferencias (no publicadas hasta la fecha) impartidas en la Universidad de California y en otros lugares.

41. Por ejemplo, la afirmación fáctica de que el mundo tiene más de 6.000 años no tiene la propiedad de poder ser probada a todas las personas inteligentes (recientemente encontré un artículo de Paul Rosenbloom –el autor del libro de lógica matemática avanzada que trabajé en mis días de estudiante

la que mencioné) *tengan status* cognitivo. Ésta es la opción más interesante abierta ante él, y hay al menos algunos filósofos y científicos sociales que están tentados por ella. Pero, ¿qué significaría creer *seriamente* una cosa así?

El hecho es que toda persona culta tiene una imagen de la Edad Media, del Imperio Romano, del Renacimiento, etc. Estas imágenes no se paran en detalles; ya que la historia ha conseguido su madurez como disciplina intelectual en el siglo XIX, han sido repetidamente revisadas, y el trabajo de los cliómetras y los historiadores sociales está conduciendo aún a más revisiones en la actualidad. Otras innovaciones en el método histórico conducirán todavía a más revisiones en el futuro. Pero, aunque mis imágenes del pasado no fueran absolutamente correctas, las considero mejor que nada. Adopto hacia ellas la actitud que Reichenbach habría recomendado. ¿Qué significaría considerar como *ficción* todo lo que creo sobre el pasado excepto el «desnudo recuerdo histórico» (*¿existe* tal cosa?)?

De nuevo, considérense las opiniones *políticas*, en la medida en que se relacionen con temas «fácticos», como opuestos a temas «morales». Cada uno de nosotros tiene muchas opiniones sobre «lo que ocurriría si» esta o aquella política, o política exterior, o económica, o política social, fuera llevada a cabo, y muchas opiniones sobre «lo que hubiera ocurrido»

---

graduado– defendiendo su derecho a creer que Dios creó el mundo hace menos de 6.000 años). Rosenbloom afirma que Dios puede haber creado el mundo en la época en que el judaísmo tradicional da como el momento de la creación, pero lleno de monumentos, testimonios escritos, gente con recuerdos falsos, etc., como en la conocida hipótesis escéptica acerca del «mundo existe desde hace cinco minutos».

si esta o aquella política *no hubiera* sido llevada a cabo. Aunque algunas de nuestras concepciones morales no dependan de esas creencias, la mayoría de nosotros somos sensibles a las consecuencias de decidir qué políticas entrañan nuestras concepciones morales; si no se tienen en absoluto creencias sobre qué acciones evitarán más probablemente la guerra, o traerán la paz a Oriente Medio, o generarán empleo, o proporcionarán refugio a los que no tienen un hogar, entonces ¿cómo se podría tener siquiera una posición política? Con todo, estos juicios todavía se resisten al tipo de acuerdo que los colocaría más allá de la controversia. Incluso cuando predecimos correctamente cuál será el resultado de una política dada, las moralejas que obtenemos del éxito de nuestra predicción son normalmente discutidas con pasión. Y es fácil discutirlas, ya que nuestras afirmaciones sobre «lo que hubiera ocurrido» si en su lugar se hubieran puesto en marcha *otras* políticas (las defendidas por nuestros oponentes políticos, en el caso en el que fuera nuestra política la que se llevara a cabo y tuviera éxito como se preveía; las que defendemos nosotros, en el caso de que fuera llevada a cabo la política oponente), son aserciones contrafácticas tan complejas que no hay manera de «confirmarlas» o «disconfirmarlas» –no hay forma de situarlas más allá de la controversia–. Nuevamente, se *podría*, supongo, adoptar la perspectiva de que todas estas concepciones no son más que ficciones; y una consecuencia podría ser que la «verdad» en tales materias ha de ser determinada imponiendo la propia voluntad; la filosofía política verdadera es aquella que tiene éxito en imponerse a sí misma, y en resistir los intentos de derrocarla.

Lo que es erróneo en las perspectivas relativistas

(aparte de su terrorífica irresponsabilidad) es que no se corresponden para nada con cómo pensamos y con cómo seguiremos pensando. Recuerdo en una ocasión a Donald Davidson haciéndome la pregunta retórica: «¿Qué entraña decir que los lenguajes intencionales son de "segunda clase"[42] si vamos a continuar usándolos?». La pregunta era una buena pregunta pragmatista. Me parece que el corazón del pragmatismo –del pragmatismo de James y Dewey, si no del de Peirce– estaba en la insistencia en la supremacía del punto de vista del agente. Si creemos que debemos adoptar un punto de vista, usar un cierto «sistema conceptual», cuando estamos comprometidos en la actividad práctica, en el más amplio sentido de «actividad práctica», entonces no debemos avanzar simultáneamente la afirmación de que ésa no es realmente «la forma como son las cosas en sí mismas». Aunque los filósofos tradicionalmente se han permitido conservar un doble conjunto de libros de esta forma, el efecto es perpetuar al menos dos errores intelectuales: conduce a degradar la noción de *creencia* (recuerden, el pragmatismo estaba inspirado por la definición de Bain de creencia –«aquello sobre lo cual una persona está preparada para actuar»–); y conduce a caer en la ficción de que hay un punto de vista del ojo de Dios que podemos imaginar útilmente. Nuestras *vidas*

42. La referencia era a la afirmación de Quine (en *Word and Objects*) de que el discurso acerca del significado (y, en el caso de los lenguajes extranjeros, incluso acerca de la verdad y de la referencia) pertenece a nuestro «sistema conceptual de segunda clase», el que usamos cuando nuestros intereses son «heurísticos» o prácticos. El sistema conceptual de «primera clase», dice Quine, es el sistema conceptual de la física. Éste es el que describe «la verdadera y última naturaleza de la realidad».

muestran que creemos que hay creencias más o menos garantizadas sobre las contingencias políticas, sobre las interpretaciones históricas, etc.

Por supuesto, mi predicción de que continuaremos hablando de esta forma, que continuaremos hablando de los análisis políticos y de las teorías históricas como más o menos garantizadas, aunque la controversia prosiga permanentemente, puede ser falseada. Podemos llegar a pensar en la historia y la política como en algo que no es más que una lucha por el poder, con la verdad como recompensa para la concepción victoriosa. Pero, entonces, nuestra cultura –todo lo que es de valor en nuestra cultura– habrá acabado.

Puede parecer que esto no es un argumento. Se dirá que solamente he señalado que la creencia en este tipo de cientifismo –en un cientifismo que mantiene *seriamente* que todo lo que no puede ser colocado más allá de la controversia no es cognitivo– tendría consecuencias desastrosas y que esto no muestra que la concepción sea falsa. Me confieso «culpable». Como atenuante, permítanme decir que he argumentado en otro lugar que las concepciones verificacionistas (y creo que la pregunta retórica que mi interlocutor me hizo era una manifestación de un tipo prefilosófico de verificacionismo) se autorrefutan.[43] Mi propósito hoy era diferente. Mi propósito consistía en romper el agarradero que cierta imagen tiene en nuestro pensamiento; la imagen de un dualismo, una división dicotómica de nuestro pensamiento en dos reinos, un reino de «hechos» que pueden establecerse más allá de la controversia, y un reino de «valores» donde estamos siempre en

43. Véase *Reason, Truth and History*, págs. 105-113.

un desacuerdo sin esperanza. Lo que espero que hayamos recordado todos nosotros –y me incluyo a mí mismo, ya que todos volvemos a deslizarnos a veces en la imagen– es la gran extensión de nuestro pensamiento que no consiste en «juicios de valor», pero que es, por otro lado, tan «controvertido» como los juicios de valor. Ninguna persona cuerda debería creer que algo es «subjetivo» meramente porque no puede ser situado más allá de la controversia.

El método científico

La imagen que he estado discutiendo es, como he dicho, una imagen prefilosófica más que una concepción sofisticada. Una forma más sofisticada de defender la dicotomía hechos-valores consiste en argumentar, con Reichenbach,[44] que las afirmaciones factuales (*incluyendo* las clases controvertidas de afirmaciones factuales que hemos estado discutiendo) pueden ser confirmadas o disconfirmadas mediante el método científico, mientras que no pueden serlo los «juicios de valor». En *Reason, Truth and History*,[45] argumenté que esta apelación al «método científico» es vacía. Mi propia concepción, para ser francos, es que no hay tal cosa como *el* método científico. El estudio de casos de teorías particulares en física, biología, etc., me ha convencido que ningún paradigma puede ajustarse a todas las diferentes investigaciones que discurren bajo el nombre de «ciencia». Pero permítanme no presuponer nada de

44. Por ejemplo, en *The Rise of Scientific Philosophy* (capítulo 17).

45. Capítulo 8, op. cit.

esto hoy. Permítanme dar por supuesto, en favor del argumento, que Reichenbach está en lo correcto o, mejor, que una concepción como la de Reichenbach está en lo correcto (ya que la propia concepción de Reichenbach del método científico, la de que puede ser reducido *todo* él a la «regla directa de inducción», no es actualmente aceptable). Concedamos que la ética no es cognitiva. Por otro lado, aún podemos creer que las teorías históricas y los contrafácticos son cognitivamente significativos, ya que (de acuerdo con Reichenbach) pueden confirmarse usando «el método científico», aunque las inferencias no sean tan buenas como se podría desear (por ello siguen siendo controvertidas). Veamos cuáles podrían ser realmente las consecuencias.

En primer lugar, veamos lo que es problemático en las inferencias usadas en historia, si se entienden simplemente como «inferencias inductivas». Consideremos un ejemplo más simple que el que utilicé antes. En su autobiografía, Collingwood incluye un capítulo fascinante acerca de su propio trabajo sobre la Bretaña Romana. El problema que intentó resolver es la reaparición de los estilos celtas antiguos en las artes decorativas después de que Bretaña dejara de ser parte del mundo romano. Una explicación podría ser que los habitantes habían continuado produciendo tal arte *durante* el período romano, pero todas las pruebas que tenemos van contra ella. De hecho, parece que los *motivos* celtas se veían como bárbaros e incluso como perversos durante el período romano. En consecuencia, Collingwood ensaya una hipótesis psicológica: (de acuerdo con esta hipótesis) el mismo hecho de que estos *motivos* fueran tabú los convirtieron en objetos de fascinación. La gente seguía describiéndolos a sus hijos (mientras que expli-

caba que eran muy, muy «perversos»), y *esto* es lo que mantuvo viva la memoria.

Es fácil ver cómo Carnap o Reichenbach podrían ver esto como un tipo de inferencia inductiva. No discutiré con quien quiera llamarlo así. Pero démonos cuenta de esto: si es una «inducción», es una inducción del tipo más problemático. Porque nuestra evidencia a favor de la generalización psicológica de que *llamar a las cosas «malvadas» causa una fascinación que nos engancha a ellas* se deriva de casos que son diferentes del caso arte-celta-en-la-Bretaña-Romana en todo punto. Con lo que estamos tratando es, en una terminología pasada de moda, con una *inducción por analogía* más que con la aplicación de una precisa (universal o estadística) ley que haya sido científicamente probada.

Cuando Carnap y yo trabajábamos juntos en lógica inductiva, en 1953-1954, el problema que él consideraba como el *más* insoluble en toda el área de la lógica inductiva era el problema de «dar un peso adecuado a la analogía». No se conoce ningún criterio para distinguir las analogías «buenas» de las «malas», y un conocido argumento de Nelson Goodman[46] muestra que queda eliminado un criterio meramente *formal*. El mismo Reichenbach admitió que no tenía ninguna receta para prevenir la aparición de contradicciones si se seguía simplemente su procedimiento formal, y propuso ir arreglando las cosas desde una base *ad hoc* según se iba operando,

46. El argumento depende de la existencia de «predicados no proyectables». Véase *Fact, Fiction and Forecast*, de Goodman, 4.ª edición, Harvard, 1983, con una introducción de H. Putnam. Para una discusión de la relevancia del resultado de Goodman, véase mi introducción a esa edición y mi *Reason, Truth and History*, 193 sigs.

escribiendo:[47] «Con respecto al requisito de consistencia, la lógica inductiva difiere intrínsecamente de la lógica deductiva; es coherente no *de facto* sino *de faciendo*, esto es, no es un *status* actual, sino una forma que ha de hacerse». Nos recuerda el comentario de John Stuart Mill (que escribía cien años antes de Reichenbach) de que no se puede hacer ciencia siguiendo servilmente las reglas de la *Lógica* de Mill (no existe un método general, Mill señalaba, que no conduzca «en conjunción con la estupidez universal» a malos resultados). En tiempos de Mill, esto no habría parecido una observación notable; pero hoy, cuando entendemos que un algoritmo adecuadamente formalizado *debe* dar los resultados que se esperan si es seguido servilmente «incluso por un idiota», como se dice, podemos ver que Mill estaba admitiendo que no había tenido éxito en *formalizar* la lógica inductiva. Si hay un sentido objetivo en el cual algunas inferencias inductivas por analogía «están garantizadas» y otras «no están garantizadas», entonces debe de haber un sentido objetivo en el cual algunos juicios acerca de lo que es «razonable» son mejores que otros, *incluso a pesar de que no podamos ofrecer un criterio general*.

Pensemos un poco más en la inferencia de Collingwood. El principio sobre el cual descansa (que llamar a las cosas «malvadas» puede hacerlas parecer fascinantes) parece admisible. Aparte de los ejemplos sexuales, ¿tenemos realmente muchos datos que apoyen esta generalización? ¿Cuáles son exactamente? ¿Quién los ha recogido? ¿Bajo qué condiciones?

No quiero sugerir en absoluto que la generalización

47. *The Theory of Probability* (traducción inglesa), 2.ª edición, pág. 450, Berkeley: University of California Press, 1949.

psicológica de Collingwood sea inadmisible. Pienso que es altamente verosímil. Pero la alta probabilidad que le adscribo no está basada en un cuerpo de «evidencia estadística» recogida acertadamente (aunque esté parcialmente basada en ejemplos que no puedo pensar, o más bien en el *sentimiento de certeza* de que podría pensar en tales ejemplos si lo intentara), sino en mi comprensión empática acerca de «cómo funciona la gente». En *Meaning and the Moral Sciences*,[48] argumenté que la *Verstehen* (entendimiento empático) *es* metodológicamente adecuada en las ciencias sociales (que es, en términos técnicos, una fuente de «probabilidad *a priori*»). A menos que admitamos que esto es así, no podremos ver en absoluto garantizado el razonamiento histórico. Dejando fuera este factor, lo que tenemos parece una analogía dudosa a partir de no-sabemos-qué-datos recogidos bajo no-sabemos-qué-condiciones. Intenten «corregir» el «defecto» formulando una «auténtica ley estadística» apoyada por «datos auténticos» y volverán al sueño de Nagel de la «historia como sociología inmadura» y de las «leyes del desarrollo histórico».

Tampoco he escogido un caso atípico desde el cual generalizar; virtualmente toda generalización histórica se apoya en analogías, y la mayoría se apoya también en una comprensión empática acerca de «cómo la gente funciona». Incluso el materialismo histórico de Marx lo hace, aunque Marx se ponía furioso cuando Stirner se lo señalaba.[49] Si se intenta,

48. Londres: Routledge y Kegan Paul, 1978.

49. Marx explica el *origen* de la explotación señalando que con la división del trabajo (intelectual y manual) surgió la oportunidad de que los trabajadores intelectuales explotaran a los trabajadores

con Ernest Nagel, simplemente *asimilar* las inferencias que hacemos en historia a las inferencias del físico, el efecto no es mostrar que la historia es, después de todo, «ciencia» auténtica, sino hacer que todo parezca ciencia *terrible*. Y lo mismo se aplica a la evidencia a favor de las inferencias que *tenemos que hacer* cuando discutimos de política y a las peculiaridades que éstas tienen, o, de igual forma, cuando pensamos sobre el carácter y disposiciones de la gente que conocemos...

Una manera de sortear la dificultad podría consistir en complementar las explicaciones formales actuales del método científico con un conjunto añadido de reglas que determinarían qué analogías son razonables, qué predicados son proyectables, qué peso deberíamos dar a nuestras intuiciones empáticas sobre cómo la gente se va a comportar bajo unas circunstancias especificadas, etc. Pero no parece existir ninguna razón para pensar que tales reglas serían de ninguna forma más simples que una descripción completa de la psicología total de un ser humano idealmente racional. (Incluso he tenido éxito al usar argumentos gödelianos[50] para de-

---

manuales. Stirner señaló que hay una premisa psicológica oculta (una que Marx no quería admitir), a saber, que *si se tiene la ocasión de explotar a otros, se los explotará*. En lugar de proponer una premisa psicológica menos desagradable, en *La Ideología alemana*, Marx simplemente afirma: «No necesito premisas psicológicas. La explotación *surge directamente* de la división del trabajo» (sic).

50. Véase mi «Reflexive Reflections», en *Epistemology, Methodology and Philosophy of Science, Essays in Honour of Carl G. Hempel on the Occasion of his 80th Birthday*, W. K. Essler, H. Putnam, y W. Stegmüller, Dordrecht: Reidel, 1985, 143-154. Este «Festschrift» fue también publicado como vol. 22, n.os 1, 2 y 3 (enero de 1985) de *Erkenntnis*, con la misma paginación.

mostrar que si la lógica inductiva de un «ser humano idealmente racional» obedece las constricciones Finetti-Shimony estándar de «coherencia», entonces incluso un ser humano idealmente racional no podría descubrir la descripción computacional completa de esa lógica inductiva.) La esperanza de tener un método formal, capaz de ser aislado a partir de nuestros juicios sobre la naturaleza del mundo, parece haberse frustrado. Y si ensanchamos la noción de un método de manera que una formalización de la psicología completa de un ser humano idealmente racional cuente como un «método», no hay razón para pensar que un «método» en este sentido deba ser independiente de los juicios del ser humano sobre metafísica, estética o cualquier otra cosa. Después de todo, la única razón para creer que el método científico no tendría presupuestos éticos (o metafísicos) era que se suponía que era un método *formal*.

El «problema epistemológico»

Si las apelaciones al «método científico» no resuelven lo que se podría llamar el problema epistemológico en ética o en cualquier otra parte, entonces ¿qué *podemos* decir sobre él? Lo que voy a decir no debería ser una sorpresa para nadie que haya seguido estas conferencias. Por ejemplo, no pienso que se pueda producir un cortocircuito en todo el tema en ética diciendo que las creencias éticas no son realmente *creencias* sino otra cosa, digamos «expresiones de una actitud» o «actos de habla distintos de decir lo que es verdadero o falso». Estos movimientos, que alguna vez estuvieron muy extendidos, eran sólo maneras de decir que los valo-

res no están «realmente ahí», esto es, que no están en las «cosas en sí» sino que son, más bien, «meras proyecciones». Seguro que esta línea de pensamiento *tiene* cierta verosimilitud –siempre la tiene, y por la misma razón–. El problema con ella no es que no «funcione», que no tenga éxito en disolver o evadirse del problema que se quiere disolver o evadir, sino que *funciona demasiado bien*. Es, por así decirlo, un solucionador general de problemas filosóficos. Uno puede, naturalmente, «librarse» de toda la problemática del *status* epistemológico de las creencias morales negando que realmente decimos algo que sea verdadero o falso cuando hacemos juicios morales, y diciendo que los valores son sólo «sentimientos que proyectamos en el mundo». Como señalé en la primera conferencia, uno puede librarse del problema de la existencia de entidades matemáticas exactamente de la misma forma, diciendo que realmente no hay cosas tales, que son objetos que meramente proyectamos en el mundo, y que afirmaciones tales como «todo número tiene un sucesor», que parecen afirmar la existencia de un número infinito de «entidades matemáticas problemáticas» no son después de todo realmente verdaderas. Simon Blackburn[51] es muy aficionado a realizar este movimiento, y no encuentra problemas en el hecho de que en su concepción la ciencia, que él afirma que realmente trata sobre las cosas tal como son en sí mismas, dependa en gran parte de premisas que no son en absoluto verdaderas sino que son meras «proyecciones». Uno puede también librarse del problema de cómo las afirmaciones *mo-*

51. Véase Blackburn, *Spreading the Word*, Oxford: Clarendon, 1984, especialmente el capítulo 6.

dales puedan ser verdaderas, es decir, cómo puede ser verdadero que ciertos estados de hechos sean «posibles», o cómo algo pueda ser «necesario». Uno puede librarse del problema de cómo las afirmaciones causales o las afirmaciones contrafácticas puedan ser verdaderas. Se puede resolver casi *cualquier* problema filosófico, o diciendo que los objetos en cuestión no existen realmente, o diciendo que las afirmaciones en cuestión no tienen realmente un valor de verdad. Si se está suficientemente desesperado o se es suficientemente atrevido, se puede decir que la misma «verdad» es una proyección (éste es, por lo que yo puedo entender, el contenido de la popular «teoría desentrecomilladora de la verdad»).[52]

Por otro lado, rechazar el punto de vista del espectador, adoptar el punto de vista del agente hacia mis propias creencias morales, y reconocer que *todas* las creencias que encuentro indispensables para la vida deben ser tratadas por mí como aserciones que son verdaderas o falsas (y las cuales creo que son verdaderas) sin una odiosa distinción entre *noumena* y *phenomena*, no es lo mismo que recaer en el realismo metafísico sobre las propias creencias morales, no más de lo que adoptar esa actitud hacia las propias creencias sobre los objetos materiales del sentido común sería recaer en un realismo metafísico sobre los objetos del sentido común, o sobre la causalidad, o sobre los objetos/modalidades matemáticas. Ello tampoco nos obliga a abandonar nuestro pluralismo o nuestro falibilismo: no

52. Véase mi «On Truth», en *How Many Questions, essays in honor of Sidney Morgenbesser*, Leigh S. Cauman *et al.* (comps.), Indianapolis: Hackett, 1983, 35-56.

se tiene por qué creer en una única *mejor* versión moral, o en una única *mejor* versión causal, o en una única *mejor* versión matemática; lo que tenemos son *mejores y peores* versiones, y en esto consiste la objetividad.

Una vez que hemos abandonado la imagen de una totalidad de objetos noumenales y propiedades a partir de las cuales nuestros diferentes esquemas conceptuales meramente realicen una u otra selección, la imagen de una masa de pastelería noumenal que nuestros esquemas conceptuales simplemente «cortan» de manera diferente, estamos obligados a reconocer con William James que la pregunta acerca de qué parte de nuestra telaraña de creencias refleja el mundo «en sí mismo» y qué parte constituye nuestra «contribución conceptual» no tiene más sentido que la pregunta:[53] «¿Anda un hombre más esencialmente con su pierna izquierda o con su pierna derecha?». El rastro de la serpiente humana está por todas partes.

En particular, nuestras creencias morales, en mi opinión, no son aproximaciones a las verdades morales del propio universo, de la misma forma que nuestras creencias científicas no son aproximaciones a la teoría científica del propio universo. Pero, no concluyo a partir de aquí que la verdad sea sólo una cuestión de lo que los compañeros de mi cultura crean. Ruth Anna Putnam[54] ha escrito que

53. La séptima Conferencia de James sobre el pragmatismo «Pragmatism and Humanism», en *Pragmatism and the Meaning of Truth*, Cambridge, Ma: Harvard University Press, 1978, págs. 115-130 (trad. cast. [en dos obras], *Pragmatismo*, Madrid, Aguilar, 1959 y *El significado de la verdad*, Madrid, Aguilar, 1958).

54. «Creating Facts and Values», *Philosophy*, vol. 60, abril de 1985.

«construimos» hechos y «construimos» valores; pero el que construyamos hechos y valores no significa que éstos sean arbitrarios o que no puedan ser mejores o peores. Ella compara la situación con la construcción de artefactos; *literalmente* construimos artefactos, y no lo hacemos de acuerdo con el anteproyecto de la propia naturaleza, ni hay siempre un diseño que se imponga a todos los diseñadores por ley natural (cuando hacemos cuchillos, no seguimos el diseño del propio universo para un cuchillo), pero no se sigue que los cuchillos que fabricamos no satisfagan necesidades reales, y ciertamente los cuchillos pueden ser mejores o peores.

Creo que sus ideas son adecuadas. Las «imágenes morales» de las que hablé en la última conferencia son creaciones humanas. La imagen moral del republicanismo cívico o del comunitarismo, la imagen moral de la fraternidad humana, la imagen moral de los humanos como hechos a semejanza de Dios, y la imagen moral kantiana del reino de los fines, o de agentes autolegisladores, son todas ellas creaciones humanas. Pero esto no quiere decir que las afirmaciones que hagamos, usando el lenguaje de una u otra de tales imágenes morales, no puedan ser correctas o equivocadas. El hecho de que sean creaciones humanas significa, sin embargo, que en principio pueden ser reemplazadas, fusionadas, combinadas, etc. –todas las operaciones sobre las «versiones» que Nelson Goodman describe en su profundo pequeño libro[55] acerca de «la construcción de mundos» se aplican también a la construcción de mundos morales.

Permítanme volver a la analogía de Ruth Anna

55. *Ways of Worldmaking*.

Putnam entre la construcción-de-valores y la construcción-de-cuchillos, o, en mi terminología, entre manufacturar imágenes morales y manufacturar artefactos. Hay una objeción obvia, que ella reconoce y discute. Alguien puede decir «sí, estás diciendo que los valores que construimos, o las imágenes morales que construimos, son mejores o peores. Pero esto supone que existen criterios para determinar lo que es mejor o peor. ¿No reaparecerá todo el problema de una epistemología o una ontología para el discurso moral cuando preguntemos sobre el *status* de estos criterios? Cualquier buen goodmaniano, o cualquier buen quineano, debería ser capaz de pensar la respuesta a esto una y otra vez. La representación que está en juego es, por supuesto, que los criterios pre-existen, es decir, preexisten los criterios para hacer inferencias inductivas, o los criterios para construir imágenes morales, o cualquier otra cosa; y las cosas que hacemos han de ser juzgadas en relación con esos cánones pre-existentes». Pero ésta es justamente la imagen que atacó Goodman en su famoso escrito sobre la inducción,[56] y que Quine ataca en su «epistemología naturalizada». Los criterios y las prácticas, han insistido siempre los pragmatistas, deben ser desarrollados juntos y ser constantemente revisados mediante un procedimiento de delicado ajuste mutuo. Los propios criterios por los cuales juzgamos y comparamos nuestras imágenes morales son creaciones tanto como las imágenes morales.

Pero, ¡un momento! Si esto es así, ¿no estamos volviendo hacia un irracionalismo completo? Ruth Anna Putnam, siguiendo la línea de John Dewey,

56. *Fact, Fiction and Forecast.*

apela en este punto a la noción de *necesidad*. Es porque hay necesidades humanas reales, y no simplemente deseos, por lo que tiene sentido distinguir entre valores mejores y peores, y, por lo mismo, entre cuchillos mejores y peores. ¿Qué son estas necesidades humanas pre-existentes y cómo se distinguen de los meros deseos? Aquí, Dewey, como Goodman, nos dice que las necesidades humanas tampoco pre-existen, que la humanidad está constantemente rediseñándose a sí misma, y que nosotros *creamos* las necesidades. De nuevo, muchos tendrán la sensación de vértigo, o peor, la sensación de caer en un pozo sin fondo. Nuestras nociones –la noción de un valor, la noción de una imagen moral, la noción de un criterio, la noción de una necesidad– están tan entrelazadas que ninguna de ellas puede proporcionar un «fundamento» para la ética. Esto, pienso, es exactamente correcto. Debemos acabar viendo que no hay posibilidad de una «fundamentación» para la ética, de la misma forma que hemos acabado viendo que no hay posibilidad de una «fundamentación» del conocimiento científico, o de cualquier otra clase de conocimiento.

Me gustaría finalizar ilustrando nuestra posición con la ayuda de una analogía, aunque para trazar la analogía debo asumir el riesgo de olvidar la cuestión que hemos estado discutiendo, la cuestión de la ética y de la «justificación» de las imágenes morales. Más aún, hay razones por las que puede parecer algo así como un abrupto cambio de tema en el mismo final de mis conferencias. Estamos obsesionados por cierta distinción culturalmente aceptada entre «ciencia» y «ética», pero estamos también obsesionados por otra distinción culturalmente aceptada, la distinción entre valores «absolutos»

e «instrumentales» –en efecto, la distinción entre evaluación e *ingeniería*–. El mismo Kant estaba atrapado por esta dicotomía cuando insistía en que todos los «imperativos» debían ser o «hipotéticos» o «categóricos». Se ha supuesto siempre que los «imperativos hipotéticos», las afirmaciones sobre lo que se debe hacer *si* se quiere conseguir un fin particular, no son problemáticos exactamente en el mismo sentido en el que los enunciados científicos son considerados no problemáticos. Mi propósito en mis comentarios finales será mostrar que esto es erróneo: que si estamos en una situación que parece problemática en «ética», estamos exactamente en la misma posición en «ingeniería», que el imperativo hipotético está en la misma situación que el categórico, que la racionalidad es una cosa igualmente difícil de «explicar» en ambos casos.

El caso que quisiera discutir es uno que Peirce usó para trazar cierta conexión entre los problemas científicos y los problemas éticos –aunque no la que yo deseo trazar–. En mi opinión, la gran contribución de Peirce descansa en su percepción de la profundidad de los problemas particulares, aunque él no tuviera éxito en construir un sistema unificado a partir de aquellas maravillosas percepciones. Una de esas grandes intuiciones geniales tiene lugar cuando Peirce discute la cuestión, ¿por qué una persona debería hacer lo que es más *probable* que funcione?

Supongan que estoy en una situación en la cual tengo que hacer X o Y y la probabilidad de éxito es muy alta si hago X y muy baja si hago Y. Podemos presentar la cuestión de Peirce de esta forma: ¿Por qué debería hacer X? ¿Por qué el hecho de que X tendrá probablemente éxito es una razón para hacerlo?

La importancia del puzzle de Peirce

Muchos filósofos dirían que la razón por la que uno debería guiarse por las probabilidades consiste en que la *frecuencia del éxito* que se disfrutaría será más alta si se hace así. Obsérvese que el caso no es uno en el cual las mismas probabilidades sean completamente inciertas; se supone que conocemos las probabilidades, y así el problema de la inducción, es decir, el problema de averiguar las probabilidades, no se plantea aquí. El asunto es que sabemos que la probabilidad de éxito es alta si se hace X, baja si se hace Y, y la pregunta es ¿por qué deberíamos hacer X? Obsérvese también que el conocimiento dado es precisamente del tipo que se supone que «justifica» el imperativo categórico «haz X si quieres tener éxito».

Es en este punto del argumento donde el genio de Peirce se manifiesta. Supóngase que soy un viejo, o que por alguna otra razón creo que no tengo muchos años de vida por delante. ¿Qué tienen que ver las creencias sobre lo que sería mi frecuencia de éxito si *fuera* a vivir mucho tiempo y estuviera inmerso en un gran número de estas situaciones con *esta* situación particular? De hecho, Peirce considera una situación en la cual la elección se da entre la «felicidad eterna» y la «aflicción perpetua». Por la misma naturaleza de esta situación, no va a *haber* ninguna «situación de juego» *posterior* con la que tenga que enfrentarse el agente racional. Específicamente, el ejemplo imaginario de Peirce es éste:[57] hay que escoger entre dos opciones. Cada opción es probabi-

57. Peirce analiza este ejemplo en «The Doctrine of Chances», pág. 69, reimpreso en *Change, Love and Logic*, Morris R. Cohen (comp.), Nueva York: Hartcourt, Brace, 1923.

lista; bajo cada opción se seleccionará una carta de un grupo bien barajado de 25 cartas, una de las cuales se marca de forma especial. El resultado depende en ambos casos de si uno consigue o no la carta así marcada. Bajo la opción A, se consigue la aflicción eterna si se saca la carta designada y la felicidad eterna si se saca cualquier otra carta, de manera que las oportunidades de felicidad eterna son veinticuatro a una; mientras que bajo la segunda opción las cosas son al revés –se consigue la felicidad eterna si se saca la carta designada y la aflicción perpetua si se saca cualquier otra carta, de manera que las oportunidades de aflicción eterna son veinticuatro a una–. (Aquellos para quienes la noción de inmortalidad sea problemática pueden sustituir la felicidad eterna y la aflicción eterna por «una muerte fácil» y «una muerte difícil», respectivamente.) *Todos* nosotros creemos que una persona racional escogería la opción A. La pregunta de Peirce es *¿por qué debería hacerlo?*

Reichenbach mantuvo que las afirmaciones de probabilidad en relación con un único caso son simplemente una *transferencia ficticia* de frecuencias relativas a largo plazo,[58] o del conocimiento de frecuencias relativas a largo plazo. Dense cuenta de que éste es aún otro ejemplo del uso de la noción de proyección. Reichenbach estaba diciendo que la misma afirmación de que Jones sólo tendrá una oportunidad de felicidad eterna entre veinticinco, *esta vez* bajo la opción B, es una «proyección». No hay ningún hecho acerca de la situación singular irrepetible que sea *el hecho de que escoger A le da a Jones veinticuatro oportunidades de felicidad eterna*

58. Reichenbach discute el caso único en *The Theory of Probability*, 372 sigs.

de cada veinticinco. (Recientemente, Stephen Leeds ha escrito un interesante artículo[59] argumentando que toda la noción de probabilidad es una proyección.)

El problema de Peirce aparece muy claramente si adoptamos la perspectiva de que la probabilidad sólo es frecuencia relativa a largo plazo. La persona en la situación conoce un hecho que es *del todo irrelevante* respecto a lo que ella debe hacer. Sabe que *si hubiera* una serie de situaciones como ésta, entonces *tendría* la felicidad eterna veinticuatro veces de cada veinticinco si escogiera la opción A todas las veces. ¡Pero una persona puede tener la felicidad eterna o la aflicción perpetua sólo una vez! Su problema no consiste en conseguir la felicidad eterna veinticuatro veces *de cada veinticinco*; su problema es obtener la felicidad eterna *esta* vez. ¿Por qué debería escoger la opción A?

La única respuesta que podemos dar es que es más probable que tenga la felicidad eterna bajo la opción A. Pero la cuestión es, recuerden, ¿por qué se debería *esperar* lo que es *probable*? Si dicen que deberían esperar lo que es *probable* porque es *posible* que ocurra esta vez, no están respondiendo a la pregunta, están únicamente repitiendo el consejo: *esperar lo que es probable*. Si dicen, «bien, es *razonable* esperar lo que es probable», bueno, en esta situación, ¿no es «razonable» sólo un sinónimo de «probable», en el sentido de Keynes-Carnap de «probabilidad lógica»? ¿No es «es razonable esperar lo que es probable que ocurra» sólo otra forma de decir «es *probable* (en el sentido lógico de probabili-

59. Véase Leeds, «Chance, Realism, Quantum Mechanics», *Journal of Philosophy*, 81 (1964), págs. 97-107.

dad) que lo que probablemente ocurrirá (en el sentido de la probabilidad como frecuencia) ocurrirá en cualquier caso individual (a menos que conozcamos algún aspecto en el cual el caso individual sea atípico)»?

Estamos forzados, entonces, a volver a la concepción según la cual una persona razonable ajusta sus expectativas de acuerdo con la probabilidad *lógica*; y esta vez, cualesquiera creencias que podamos tener sobre cómo esto nos guiará bien a largo plazo se consideran irrelevantes para el problema. Que existe tal cosa como la «probabilidad lógica», que se corresponde con la frecuencia en una serie larga (si *hubiera* una serie larga), y que una persona razonable ajusta sus creencias a *ella*, pasan a ser sólo hechos lógicos (*léase*: metafísicos) últimos.

La propia solución de Peirce a este problema es una de las fuentes de inspiración para las concepciones de Apel y Habermas que mencioné en la última conferencia. Según Peirce, *sólo* se puede ser racional si uno se *identifica* a sí mismo psicológicamente con toda una continua (de hecho, potencialmente infinita) comunidad de investigadores. Es únicamente porque me preocupo sobre lo que *podría* ocurrir a los demás en situaciones similares por lo que hago lo que tiene mayores *oportunidades* en mi propia situación. En la perspectiva de Peirce, mi creencia de que de alguna forma, en esta situación *irrepetible*, es más probable que yo experimente la felicidad eterna que la aflicción eterna es fundamentalmente, entonces, sólo lo que Reichenbach decía que era, una transferencia ficticia. Lo que es verdad, y no una ficción o una proyección, sin embargo, es que mis compañeros, los miembros de la comunidad con la cual me identifico, tendrán la fe-

licidad eterna veinticuatro veces de cada veinticinco si siguen esta estrategia; o, más en general, aunque no se repita nunca esta situación particular, si en todos los diferentes casos *no relacionados* de esta clase, o de cualquier otra clase que ellos mismos encuentren, siempre siguen las probabilidades, entonces a largo plazo experimentarán más éxitos y menos fracasos.

Pero, ¿puede ser realmente el que yo sea *altruista* la razón por la que escogería la opción A? Tal vez yo lo sea, aunque ¿no resulta obvio que escogería la opción A en primer lugar y ante todo porque ello evitaría la aflicción perpetua *en mi propio caso*? El argumento de Peirce consiste en que debo escoger la opción A por lo que se podrían describir como razones para un «utilitarismo de la regla»: al escoger esta opción estoy apoyando, y ayudando a perpetuar, una regla que beneficiará a la humanidad (o a la comunidad de investigadores racionales) *a largo plazo*. ¿Es realmente esto lo que tengo en mente cuando con lo que me estoy enfrentando es con la *tortura* (aflicción perpetua)? Francamente, no lo es. Si las únicas razones permitidas han de ser en términos de «lo que ocurrirá a largo plazo si», no puedo dar una razón para hacer lo que haría en este caso. Y esto muestra que incluso en el tipo de problema *medios-fin*, debo recaer en la razón sin formalizar.

Actualmente, mucha gente[60] piensa que la única razón para ser *razonable* radica en que se llegará a la verdad en la teoría, y al éxito en la acción, *más a me-*

60. Para una discusión de esta clase de «utilitarismo epistémico», véase el discurso presidencial de Roderick Firth «Epistemic Merit, Intrinsic and Instrumental», en *Proceedings and Addresses of the American Philosophical Society*, septiembre de 1981, vol. 55, n. 1, 5-23.

nudo si se es razonable. Alguna gente[61] incluso se ha propuesto reemplazar la noción de un método «razonable» por la noción de un método confiable: un método tal que, como un asunto de hecho, conduce a resultados satisfactorios con una alta frecuencia relativa. Dense cuenta de que (si están de acuerdo conmigo en no considerar creíble la propia solución de Peirce) estas perspectivas no son de ninguna ayuda frente al problema de Peirce. Si mi *única* razón para creer que debería ser razonable fueran mis creencias sobre lo que ocurrirá *a largo plazo* si actúo o creo de manera razonable, entonces no tendría en absoluto *ninguna* razón (aparte de la inverosímil razón del altruismo) para pensar que es mejor ser razonable en un caso singular irrepetible como el descrito. De hecho, según me acerque al final de mi vida, y me encuentre a mí mismo incapaz de hacer más «apuestas», entonces mis razones para hacer lo que es razonable o esperar lo que es razonable disminuirán bruscamente desde este punto de vista. El hecho es que tenemos cierto tipo de obligación *primitiva* y *no derivada* de ser razonables; ciertamente no es una «obligación moral» o una «obligación ética», pero no obstante es una obligación muy real de ser razonables que, en oposición a Peirce, *no* es reducible a mis expectativas a largo plazo y mi interés en el bienestar de los demás o en mi propio bienestar en otras ocasiones. Ciertamente, *también* creo que ser razonable funcionará mejor a largo plazo para la gente; pero cuando la cuestión es *¿por qué esperar que, en este caso* irrepetible, *lo que es extremadamente proba-*

61. Para una versión sofisticada de esta concepción, véase de Alvin I. Goldman «What is Justified Belief?», en George Pappas (comp.), *Justification and Knowledge*, Dordrecht: Londres, 1979.

ble que ocurra ocurrirá?, en este punto tengo que decir con Wittgenstein: «Aquí es donde mi pala se dobla. Esto es lo que hago, esto es lo que digo».

La razón que tengo para traer esto a colación cuando la cuestión era qué hacer en relación al fenómeno del «pozo sin fondo» en ética, con nuestra preocupación por la falta de fundamentación de la ética, es que en el caso que se acaba de describir –un caso que tiene que ver con la razonabilidad sobre «medios y fines», más que con la ética– mi situación epistémica es exactamente la misma. Creo, y creo que estoy justificado en creerlo, que «actuar sobre la base de las probabilidades» es lo único racional que cabe hacer, y que se debe hacer lo racional incluso en situaciones irrepetibles. En el caso ético, creo, y creo que estoy justificado en creerlo, que una persona con sentido de la fraternidad humana es mejor que una persona que carece de sentido de la fraternidad humana. Una persona que es capaz de pensar por sí misma sobre cómo vivir es mejor que una persona que ha perdido, o nunca ha desarrollado, la capacidad de pensar por sí misma cómo vivir; pero, ya se plantee la cuestión sobre la probabilidad en el caso singular o sobre la ética, no *sé como sé* estas cosas. Son casos en los cuales encuentro que tengo que decir: «He llegado a un lecho de roca y aquí es donde mi pala se dobla».[62]

62. *Philosophical Investigations*, Oxford: Blackwell, 1953, sec. 217. Que Wittgenstein use aquí la primera persona –donde mi pala se dobla– es muy importante; con todo, muchos intérpretes intentan ver su filosofía como si ésta simplemente respetara alguna «forma de vida» determinada por una comunidad. Sobre esto, véase también la discusión de Stanley Cavell en *The Claim of Reason*, especialmente la parte uno, capítulo V, «The Natural and The Conventional».

Reconocer que hay ciertos lugares donde a uno se le dobla la pala; reconocer, con Wittgenstein, que hay lugares donde nuestras explicaciones se acaban, no es decir que algún lugar particular está *permanentemente* destinado a ser «un lecho de roca», o que alguna creencia particular es para siempre inmune a la crítica. Aquí es donde mi pala se dobla *ahora*. Aquí es donde mis justificaciones y explicaciones acaban *ahora*. Reconocer que un ser humano leal es mejor que un ser humano desleal, que una persona capaz de *philia* es mejor que una persona incapaz de *philia*, que una persona capaz de sentido de la comunidad, de ciudadanía en una *polis*, es mejor que una persona que es incapaz de sentido de la comunidad o de ciudadanía en una *polis*, etc., no es decir que cualquiera de estos valores o cualquiera de estas imágenes morales, que pueden subyacer a estos valores y organizarlos, sea última, en el sentido de ser exclusiva o exhaustivamente correcta. Nuestras imágenes morales están en un proceso de desarrollo y reforma. Pero esto quiere decir que en cada nivel de este desarrollo y reforma habrá lugares, muchos lugares, en los cuales tengamos que decir: «Aquí es donde mi pala se dobla».

Nada de esto va en contra de la idea de que sea posible la crítica racional de una visión moral. Una visión moral puede contradecir, por ejemplo, lo que desde otra base sabemos o pensamos que es racional creer, sea ello algo lógico, metafísico o empírico. Pero no podemos esperar durante más tiempo que esta clase de críticas acabe dejando sólo una *única* visión moral intacta. Y al final siempre hay un punto en el cual se tiene que decir: «Aquí es donde mi pala se dobla».

Esto *no* es relativismo feyerabendiano. *Es* un re-

chazo del proyecto de la Epistemología con una «E» mayúscula. En la primera lectura, intenté mostrar que rechazar el proyecto de la ontología –de una descripción de las cosas tal como son «aparte de nuestros sistemas conceptuales»– no pone fin a todas las cuestiones interesantes sobre el lenguaje y el pensamiento; más bien llama la atención sobre fenómenos que hemos estado soslayando (cuando no realmente ignorándolos), por ejemplo, el fenómeno que llamé relatividad conceptual. Y, de modo similar, rechazar el proyecto de la Epistemología con una «E» mayúscula –el proyecto de un método universal para decir quién tiene «la razón de su lado» no importa de *qué* disputa se trate– no pondrá fin a todas las cuestiones interesantes sobre el conocimiento en ciencia y en ética; en lugar de ello, dirigirá nuestra atención a otros fenómenos que hemos estado intentando ignorar (por ejemplo, las analogías entre los problemas epistemológicos –dense cuenta de la «e» minúscula– conectados con la ciencia y los problemas epistemológicos conectados con los valores). Sobre todo, espero que pueda reorientar la energía filosófica hacia una de sus tareas tradicionales –la única tarea que la filosofía no debería nunca abandonar–, la tarea de ofrecer imágenes de la situación humana en el mundo discutibles, importantes y llenas de significado.